



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 111157

TO: Nirmal Basi
Location: cm1/10e17
Art Unit: 1646
Friday, January 02, 2004

Case Serial Number: 09/719485

From: Mary Jane Ruhl
Location: Biotech-Chem Library
CM1-6A06
Phone: 605-1155

maryjane.ruhl@uspto.gov

Search Notes

Examiner Basi,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
CM-1, Rm. 6-A-06
605-1155

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 31, 2003, 18:55:58 ; Search time 4719.89 Seconds
(without alignments)
10739.032 Million cell updates/sec
Title: US-09-719-485-2
Perfect score: 1239
Sequence: 1 atggcgagccctggaacgg.....acgtgaagacatgggataa 1239
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sta.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1239	100.0	1239	6	AX154589 Sequence
2	1239	100.0	1239	6	AX549187 Sequence
3	1239	100.0	1239	6	AX572965 Sequence
4	901	72.7	1161	6	AX154591 Sequence
5	901	72.7	2040	9	AF034632 Homo sapi
6	901	72.7	163284	9	AL137000 Human DNA
7	901	72.7	341560	2	AL596304 Homo sapi
8	901	72.7	349980	6	AX711879 Sequence
9	901	72.7	349980	6	AX739961 Sequence
10	592.2	47.8	692	9	HS3339407 Homo sapi
11	572.2	46.2	692	9	HS3442408 Homo sapi
12	557.4	45.0	615	9	HS3339459 Homo sapi
13	424.8	34.3	701	9	HS3326768 Homo sapi
c 14	307	24.8	192116	2	AC109158 Mus muscu
15	283	22.8	283	6	AR168467 Sequence
16	283	22.8	283	6	AR182284 Sequence
17	283	22.8	283	6	AR303910 Sequence
18	283	22.8	283	6	BD056681 Galanin r
19	283	22.8	283	6	BD064652 Human gal
20	283	22.8	283	6	BD064664 Mouse gal
21	266.8	21.5	573	9	HS3441295 Homo sapi
22	244	19.7	1751	5	AB095995 Gallus ga
23	242.6	19.6	1351	5	AB095997 Gallus ga
24	242.6	19.6	1703	5	AB095996 Gallus ga
25	242.6	19.6	4121	5	AB095994 Gallus ga
26	234.8	19.0	1063	6	AR156351 Sequence
27	234.8	19.0	1101	4	SSU60178 Sus scrofa
28	233.2	18.8	870	4	SSU60180 Sus scrofa
29	230	18.6	1029	6	AR156352 Sequence
30	230	18.6	110000	2	AC095360 2 Continuation (3 of
c 31	230	18.6	183099	2	AC117349 Rattus no
32	230	18.6	195131	2	AC108400 Mus muscu
33	230	18.6	227035	2	AC116585 Mus muscu
c 34	230	18.6	228796	2	AC095173 Rattus no
c 35	230	18.6	250754	2	AC130771 Rattus no
36	229.8	18.5	1092	6	AR156357 Sequence
37	229.8	18.5	1092	6	BD094703 GHSR liiga
38	229.8	18.5	1092	6	ES4510 Novel poly
39	229.8	18.5	1350	10	AB001982 Rattus no
40	229	18.5	250	6	AR269796 Sequence
41	228.8	18.5	870	6	AX548854 Sequence
42	228.8	18.5	870	9	HSU60181 Human growt
43	228.8	18.5	1122	6	AR156354 Sequence
44	228.8	18.5	6787	9	AF369786 Homo sapi
c 45	228.8	18.5	145593	9	AC069523 Homo sapi

ALIGNMENTS

RESULT 1
AX154589
LOCUS AX154589 1239 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 9 from Patent WO0138155.
ACCESSION AX154589
VERSION AX154589.1 GI:14536175
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Sheppard,P.O., Jaspers,S.R., Deisher,T.A. and Bishop,P.D.
TITLE Method of forming a peptide-receptor complex with zaig33 and
therapeutic use thereof


```
Qy 61 GCGCTGCGGCTTGGAGAGCGCGCTGCTCGCCCTTTCCTTGGGGGCGCTGGTGCG 120
Db 61 GCGCTGCGGCTTGGAGAGCGCGCTGCTCGCCCTTTCCTTGGGGGCGCTGGTGCG 120
Qy 121 GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCG 240
Db 181 ATGCTGATCGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCG 240
Qy 241 GCGCTGTCGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GCGCTGTCGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 301 TCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GTGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGGACCCGCGCATCTCGTAGTCCCGGCGCTCAATGGCACCGCGCGGATCGCTCTCG 600
Db 541 CAGGACCCGCGCATCTCGTAGTCCCGGCGCTCAATGGCACCGCGCGGATCGCTCTCG 600
Qy 601 CCTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CCTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GGGCCGAGACCGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GGGCCGAGACCGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGCTCATCGGGGCTCATCGGGGCTCAT 840
Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGCTCATCGGGGCTCATCGGGGCTCAT 840
Qy 841 CTGGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 CTGGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 TACATAACACGGAAGATTCGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TACATAACACGGAAGATTCGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCCATCTCTCAACCTCTCTCAACCT 1080
Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCCATCTCTCAACCTCTCTCAACCT 1080
Qy 1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1141 CACAGAGCAGGAGACATGCGGGGGAAGTTGACGGGGAACATGAGGAGACACCGTGCGG 1200
```

```
Db 1141 CACAGAGCAGGAGACATGCGGGGGAAGTTGACGGGGAACATGAGGAGACACCGTGCGG 1200
Qy 1201 TACACCGAGACAGCGCTAAACGCTGAAGACGATGGGATAA 1239
Db 1201 TACACCGAGACAGCGCTAAACGCTGAAGACGATGGGATAA 1239

RESULT 3
AX572965 1239 bp DNA linear PAT 29-NOV-2002
LOCUS Sequence 1 from Patent WO02057791.
DEFINITION AX572965
ACCESSION AX572965
VERSION AX572965.1 GI:26005013
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Brown,J.P., Burmer,G.C., Roush,C.L. and Kulander,B.G.
TITLE Diagnostic and therapeutic compositions and methods related to gpr
38
JOURNAL Patent: WO 02057791-A 1 25-JUL-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1239
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 172 a 431 c 397 g 239 t
ORIGIN

Query Match 100.0%; Score 1239; DB 6; Length 1239;
Best Local Similarity 100.0%; Pred. No. 1.8e-155;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGACGACCCCTGGAAACGCGACGACGCGCCCGAGGGGCGCGGAGCGCGCGCGCC 60
Db 1 ATGGGACGACCCCTGGAAACGCGACGACGCGCCCGAGGGGCGCGGAGCGCGCGCGCC 60
Qy 61 GCGCTGCGGCTTTCGAGACGAGCGCGCTGCTCGCCCTTTCCTTGGGGGCGCTGGTGCG 120
Db 61 GCGCTGCGGCTTTCGAGACGAGCGCGCTGCTCGCCCTTTCCTTGGGGGCGCTGGTGCG 120
Qy 121 GTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCG 240
Db 181 ATGCTGATCGGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCG 240
Qy 241 GCGCTGTCGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GCGCTGTCGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 301 TCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TGCACCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GTGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
```

```
Qy 541 CAGGACCCCGGCTATCCGTAGTCCCGGCTCAATGACACCGCGGATCGCTCTCG 600
Db 541 CAGGACCCCGGCTATCCGTAGTCCCGGCTCAATGACACCGCGGATCGCTCTCG 600
Qy 601 CCTCTCGCTCGTCCCGGCTCTCTGCTCTGCGGGGGGACCGCGCTCCCGCGCTCG 660
Db 601 CCTCTCGCTCGTCCCGGCTCTCTGCTCTGCGGGGGGACCGCGCTCCCGCGCTCG 660
Qy 661 GGGCCGAGACCGGAGGCGCGGCTGTTACGCGCGGAATCCCGCGGAGCCCGCGG 720
Db 661 GGGCCGAGACCGGAGGCGCGGCTGTTACGCGCGGAATCCCGCGGAGCCCGCGG 720
Qy 721 CAGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGCTGTGGAGAGCGCGGCGG 840
Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGCTGTGGAGAGCGCGGCGG 840
Qy 841 CTGGAGGCGCGGCTCGGGGAGAGAGGCGGAGAGCGGAGAGCGGAGAGCGGAG 900
Db 841 CTGGAGGCGCGGCTCGGGGAGAGAGGCGGAGAGCGGAGAGCGGAGAGCGGAG 900
Qy 901 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 961 TACATAACAGGAGATTCGGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TACATAACAGGAGATTCGGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 1021 CTGCAACTTTTCTATCTGAGCGCATCTCAACCAATCTCTCAACCAATCTCTCA 1080
Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTCAACCAATCTCTCAACCAATCTCTCA 1080
Qy 1081 AAGTACAGAGCGGCGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 AAGTACAGAGCGGCGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1141 CACAGAGAGGAGACACTGCGGGGAGAGTTGCGGGGAGACACTGAGGAGACACG 1200
Db 1141 CACAGAGAGGAGACACTGCGGGGAGAGTTGCGGGGAGACACTGAGGAGACACG 1200
Qy 1201 TACACCGAGACAGCGCTAAGTGAAGCAGATGGATAA 1239
Db 1201 TACACCGAGACAGCGCTAAGTGAAGCAGATGGATAA 1239
```

```
RESULT 4
AX154591
LOCUS AX154591 1161 bp DNA linear PAT 23-JUN-2001
DEFINITION Sequence 11 from Patent WO0138355.
ACCESSION AX154591
VERSION AX154591.1 GI:14536177
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sheppard, P.O., Jaspers, S.R., Deisher, T.A. and Bishop, P.D.
TITLE Method of forming a peptide-receptor complex with zsig33 and
therapeutic use thereof
JOURNAL Patent: WO 0138355-A 11 31-MAY-2001;
ZymoGenetics, Inc. (US)
FEATURES
source
1..1161
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
1..1161
/note="unnamed protein product"
```

```
BASE COUNT 126 a 436 c 372 g 227 t
ORIGIN
Query Match 72.7%; Score 901; DB 6; Length 1161;
Best Local Similarity 100.0%; Pred. No. 1.5e-110; Indels 0; Gaps 0;
Matches 901; Conservative 0; Mismatches 0;
Qy 1 ATGGGAGCGCCCTGGAAACGGCAGCGAGCGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60
Db 1 ATGGGAGCGCCCTGGAAACGGCAGCGAGCGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60
Qy 61 GCGCTGCGCGCTTGGCAGAGAGCGCGCTGCTCGGCCCTTTCCTTGGGGGGCGCTGGTGG 120
Db 61 GCGCTGCGCGCTTGGCAGAGAGCGCGCTGCTCGGCCCTTTCCTTGGGGGGCGCTGGTGG 120
Qy 121 GTGACCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTGACCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGGGCTACCGGGAGATCGGGAGACCAACCAACCAACCAACCAACCA 240
Db 181 ATGCTGATCGGGGGCTACCGGGAGATCGGGAGACCAACCAACCAACCAACCAACCA 240
Qy 241 GCGGTGTCGACCTACTCATCTGCTGCGGGTGGCGGTTCGACCTGACCGCTCTGGCGG 300
Db 241 GCGGTGTCGACCTACTCATCTGCTGCGGGTGGCGGTTCGACCTGACCGCTCTGGCGG 300
Qy 301 TCGCGCGCTGGGTGTTTCGGGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGCGCGCTGGGTGTTTCGGGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGACCTTACGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 TGACCTTACGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 421 TCGCGCGCTGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCGCGCGCTGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GTGCTGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGGACCCCGGCTATCTCCGTAGTCCCGGGGCTCAATGGGACCGCGGAGATCGCTCTCG 600
Db 541 CAGGACCCCGGCTATCTCCGTAGTCCCGGGGCTCAATGGGACCGCGGAGATCGCTCTCG 600
Qy 601 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GGGCCGAGAGCGGAGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GGGCCGAGAGCGGAGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 CAGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGCATCTCTAAGGGCTCATCGGGGAGAGCTGTGGAGAGCGCGGCGGCGG 840
Db 781 CTGTGCTCAGCATCTCTAAGGGCTCATCGGGGAGAGCTGTGGAGAGCGCGGCGGCGG 840
```

```
/codon_start=1
/protein_id="CAC42612.1"
/db_xref="GI:14536178"
/translation="MGSPWNGSDGPEGAREPPHPALPPCDERRCSPPPLGALVPTAV
CLCLFVVGSGNVVTVMLIGRYRDMRTTINLYIGSMVAVDLLILGLGDFVFDIYLRWR
PWGFPYLLCRSLYVGEGETYATLLHMTALSVRYLAICRPLARLVLTFRVRRALRIA
VLWAVALLSAGAPFLVGVQDPGIVSVVPLNGTARIASPLASSPPLMWSRAPPPSP
PSPETAEEAALFSRECRPSAOLGALRWMLVTTAVFFPLFLCLSLYGLIGRELWS
SRPLRGPAASGRGHRGTIVRLRWKSRGSKDACLQSAAPGPTAOTGLPLPLLAQLW
APLAPPPISIPASTRRGGSGIYNLLVALPWRQNHLLKHGRFADDDVLLSVL"
```

```
Qy 841 CTCGAGAGCCCGCGCCCTCGGGCGGAGAGAGGCCACCGGACACCGTCCGCGTCTGT 900
Db 841 CTCGAGAGCCCGCGCCCTCGGGCGGAGAGAGGCCACCGGACACCGTCCGCGTCTGT 900
Qy 901 C 901
Db 901 C 901

RESULT 5
AF034632 2040 bp DNA linear PRI 13-JUL-1998
LOCUS Homo sapiens orphan G protein-coupled receptor (GPR38) gene,
DEFINITION complete cds.
ACCESSION AF034632
VERSION AF034632.1 GI:2654158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2040)
AUTHORS McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighner,S.D.,
Hreniuk,D.L., Smith,R.G., Howard,A.D. and Van der Ploeg,I.H.
TITLE Cloning and characterization of two human G protein-coupled
secretagogue and neurotensin receptors
JOURNAL Genomics 46 (3), 426-434 (1997)
MEDLINE 98110578
PUBMED 9441746
REFERENCE 2 (bases 1 to 2040)
AUTHORS McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighner,S.D.,
Hreniuk,D.L., Smith,R.G., Van Der Ploeg,I.H.T. and Howard,A.D.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co.,
Inc., PO Box 2000, Rahway, NJ 07065, USA
FEATURES
source
1..2040
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="13q14-21"
<1..>2040
/gene="GPR38"
join(<1..901,1703..>2040)
/gene="GPR38"
join(1..901,1703..2040)
/note="Orphan G protein-coupled receptor related to growth
hormone secretagogue and neurotensin receptors"
/codon_start=1
/product="GPR38"
/protein_id="AAC26081.1"
/db_xref="GI:2654159"
/translation="MGSPWNGSDGEGAREPPALPCDERRCSPPFLGALVPVAV
CLCLFVGVSGNVVTMLIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSR
PWFGPLLCLRLSLVGECTYATLLHMTALSVERYLAICRLPLARVLVTRRRVALLA
VIMAVALLUSAGFFLVLVGEQDPGISVVGUNGRTARIASSPLASSPPLWLSRAPPPSP
PSGPETAEAAALFIRECPSPQAQLGALVMLVVTAYFFLPLCLSLYLGLIGRELWS
SRRLRGPAASGRGRHQTVRLVLLVLAFLIICWLPFHVGRIIVINTEDSRMYFSQ
YFNITVALQLFSLASINPILNLSIKYRAAFAKLLARKSRPRGHRSDRTAGEVAG
DTGDDTVGYTETSANVTMG"
BASE COUNT 347 a 631 c 598 g 464 t
ORIGIN
Query Match 72.7%; Score 901; DB 9; Length 2040;
Best Local Similarity 100.0%; Pred. No. 1.3e-110;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGACGCCCTTGAAACGGAGACCGGCCCGAGGGCGCGGAGACCGCGTGGCCCC 60
Db 1 ATGGGACGCCCTTGAAACGGAGACCGGCCCGAGGGCGCGGAGACCGCGTGGCCCC 60
```

```
Qy 61 GCGCTGCGCGCTTTCGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGTGGTCCG 120
Db 61 GCGCTGCGCGCTTTCGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGTGGTCCG 120
Qy 121 GTGACCGCTGTGTGCTGTGCTGCTGTTGCTGTGGGGGTGAGCGGCAACGTTGGTACCGTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGTTGCTGTGGGGGTGAGCGGCAACGTTGGTACCGTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGGACATCGGACACCAACCTTGTACTTGGGAGGATG 240
Db 181 ATGCTGATCGGGCGCTACCGGGACATCGGACACCAACCTTGTACTTGGGAGGATG 240
Qy 241 GCGGTGTCGACCTACTCATCTGCTCGGGGTGCGCTTGCACCTGTATCCGCTCTGCGGC 300
Db 241 GCGGTGTCGACCTACTCATCTGCTCGGGGTGCGCTTGCACCTGTATCCGCTCTGCGGC 300
Qy 301 TCGGGCGCTGGGTGTTGGGGCGCGCTGCTTGGCGCGCTTCCCTTACGTGGGGGAGGGC 360
Db 301 TCGGGCGCTGGGTGTTGGGGCGCGCTGCTTGGCGCGCTTCCCTTACGTGGGGGAGGGC 360
Qy 361 TGCACTTACGCCACGCTGTCACATGACCGCGCTCAGCGTTCGAGCGCTACCTGGCCATC 420
Db 361 TGCACTTACGCCACGCTGTCACATGACCGCGCTCAGCGTTCGAGCGCTACCTGGCCATC 420
Qy 421 TGGCGCGCGCTCGCGCGCGCGCTGTTGGTCAACCGCGCGCGCTTGGTCAACCGCGCGCT 480
Db 421 TGGCGCGCGCTCGCGCGCGCGCTGTTGGTCAACCGCGCGCGCTTGGTCAACCGCGCGCT 480
Qy 481 GTGCTCTGGGCGGTGGCGCTGCTCTTGGCGGTGCTTCTTGTTCCTGGGGCGTGGAG 540
Db 481 GTGCTCTGGGCGGTGGCGCTGCTCTTGGCGGTGCTTCTTGTTCCTGGGGCGTGGAG 540
Qy 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGCGACCGCGCGGATCGCTCTCG 600
Db 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGCGACCGCGCGGATCGCTCTCG 600
Qy 601 CCTCTCGCTCGTCCGCGCGCTCTCTGGCTCTCGGGGGCGGACACCGCGCTTCCCGCGCTG 660
Db 601 CCTCTCGCTCGTCCGCGCGCTCTCTGGCTCTCGGGGGCGGACACCGCGCTTCCCGCGCTG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTCAGCGCGGAAATGCGGCGGAGCGCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTCAGCGCGGAAATGCGGCGGAGCGCGCG 720
Qy 721 CAGCTGGGGCGGTGCGCTGTGCTGTGGGTACACCGCTACTTCTTCTGCTGCCCTTT 780
Db 721 CAGCTGGGGCGGTGCGCTGTGCTGTGGGTACACCGCTACTTCTTCTGCTGCCCTTT 780
Qy 781 CTGTGCTTACGATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCGCG 840
Db 781 CTGTGCTTACGATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCGCG 840
Qy 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGGCCACCGGAGACCGTCCGCGTCTGT 900
Db 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGGCCACCGGAGACCGTCCGCGTCTGT 900
Qy 901 C 901
Db 901 C 901

RESULT 6
AL137000 163284 bp DNA linear PRI 18-DEC-2000
LOCUS Human DNA sequence from clone RP11-203116 on chromosome 13 Contains
DEFINITION the gene for KIAA0970 protein, COX7CP1 (cytochrome c oxidase
subunit VIIC pseudogene 1), a novel pseudogene, the GPR38 (G
protein-coupled receptor 38) gene, ESTs, STSs, GSSs and a CpG
island, complete sequence.
ACCESSION AL137000
VERSION AL137000.6 GI:9944121
KEYWORDS HTG; COX7CP1; CpG island; cytochrome c oxidase; G protein-coupled
```

receptor; GPR38; KIAA0970.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 163284)
 REFERENCE Wall, M.
 AUTHORS Direct Submission
 TITLE Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
 JOURNAL CB10 15A, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 COMMENT On Aug 29, 2000 this sequence version replaced gi:9926419.
 During sequence assembly data is compared from overlapping clones.
 where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 13, constructed by the Sanger Centre Chromosome 13
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr13>
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-203116 It may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-452110 is at 1 in this sequence.
 The true right end of clone RP11-103318 is at 163284 in this
 sequence. This sequence has been finished according to sequence map
 criteria as follows. An attempt is made to resolve all sequencing
 problems, such as compressions and repeats, but not necessarily
 within known annotated repeat sequence elements. Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key. RP11-203116 is from the library RPCI-11.1 constructed
 by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6.
 FEATURES
 source
 1..163284
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-203116"
 /clone_lib="RPCI-11.1"
 3..121
 /note="MER20 repeat: matches 4..126 of consensus"
 repeat_region 152..261
 /note="L2 repeat: matches 20..137 of consensus"
 repeat_region 262..554
 /note="AluX repeat: matches 1..293 of consensus"
 repeat_region 555..698
 /note="MIR repeat: matches 137..255 of consensus"
 repeat_region 1012..1291
 /note="AluJ repeat: matches 7..291 of consensus"
 repeat_region 1320..1425
 /note="MER94 repeat: matches 4..114 of consensus"
 repeat_region 1591..1804
 /note="L1P8 repeat: matches 5945..6158 of consensus"
 repeat_region 3240..3621
 /note="L2 repeat: matches 2310..2691 of consensus"
 repeat_region 3824..4106
 /note="AluJ repeat: matches 1..301 of consensus"
 repeat_region 4975..5042
 /note="L2 repeat: matches 2639..2705 of consensus"
 repeat_region 6113..6408
 /note="AluX repeat: matches 1..295 of consensus"
 repeat_region 7646..7732
 /note="AluSc repeat: matches 1..85 of consensus"
 repeat_region 9887..10191
 /note="AluX repeat: matches 1..306 of consensus"
 repeat_region 12201..12247
 /note="L2 repeat: matches 2704..2750 of consensus"
 repeat_region 12472..12738
 /note="L2 repeat: matches 1304..1530 of consensus"
 repeat_region 12739..12874
 /note="FLAM C repeat: matches 2..133 of consensus"
 repeat_region 12875..12887
 /note="L2 repeat: matches 1530..1542 of consensus"
 repeat_region 12922..13227
 /note="L1MB3 repeat: matches 5871..6183 of consensus"
 repeat_region 13303..13348
 /note="MTR1E repeat: matches 1..46 of consensus"
 repeat_region 13533..13623
 /note="MTR1E repeat: matches 223..315 of consensus"
 repeat_region 13657..13865
 /note="MTR1E repeat: matches 307..533 of consensus"
 repeat_region 13890..14150
 /note="MTR1J repeat: matches 195..456 of consensus"
 repeat_region 14151..14450
 /note="AluSg repeat: matches 1..307 of consensus"
 repeat_region 14501..14729
 /note="L1M4C repeat: matches 588..813 of consensus"
 repeat_region 14730..15006
 /note="AluJ repeat: matches 1..285 of consensus"
 repeat_region 15007..15361
 /note="L1M4C repeat: matches 813..1178 of consensus"
 repeat_region 15389..15437
 /note="L1M4C repeat: matches 1159..1207 of consensus"
 repeat_region 15438..16412
 /note="HERVL repeat: matches 2891..3879 of consensus"
 repeat_region 16413..16616
 /note="MTR2B repeat: matches 1..216 of consensus"
 repeat_region 17092..17141
 /note="MER5B repeat: matches 27..79 of consensus"
 repeat_region 17416..17647
 /note="MIR repeat: matches 17..261 of consensus"
 repeat_region 19402..19499
 /note="HY3 repeat: matches 2..99 of consensus"
 repeat_region 20730..21163
 /note="L2 repeat: matches 2168..2609 of consensus"
 repeat_region 25800..25825
 /note="13 copies 2 mer tt 92% conserved"
 repeat_region 26041..26131
 /note="FLAM C repeat: matches 17..107 of consensus"
 repeat_region 26144..26249
 /note="53 copies 2 mer tt 64% conserved"
 repeat_region 27029..27259
 /note="MIR repeat: matches 2..251 of consensus"
 repeat_region 27422..27788
 /note="L2 repeat: matches 1879..2259 of consensus"
 repeat_region 27863..28239
 /note="L2 repeat: matches 1278..1734 of consensus"
 repeat_region 28255..28571
 /note="AluY repeat: matches 1..310 of consensus"
 repeat_region 28729..29128
 /note="L1MB3 repeat: matches 4905..5326 of consensus"
 repeat_region 29170..29413
 /note="L1MB3 repeat: matches 5461..5715 of consensus"
 repeat_region 29448..29681
 /note="L2 repeat: matches 1594..1828 of consensus"
 repeat_region 29683..29886
 /note="L1MB repeat: matches 5527..5734 of consensus"
 repeat_region 29888..30323
 /note="L2 repeat: matches 2129..2596 of consensus"
 repeat_region 30324..30620
 /note="AluJb repeat: matches 1..299 of consensus"
 repeat_region 30621..30780
 /note="L2 repeat: matches 2596..2749 of consensus"
 repeat_region 30784..30813
 /note="15 copies 2 mer aa 86% conserved"
 repeat_region 30814..30947

Db 38856 CTGCGAGGCCGCGCCCTCTGGGGCGGAGAGGCCACCGAGACCGTCCGCGTCTGTG 38915

QY 901 C 901

Db 38916 C 38916

RESULT 10
HSA339407
LOCUS
DEFINITION Homo sapiens genomic sequence surrounding NOT1 site, clone
NR5-1B14C.
VERSION AJ339407
KEYWORDS AJ339407.1 GI:15983825
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M., Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V., Leviscky, V.G., Kotchanov, N.A., Protopopov, A.I., Kashuba, V.I., Kissilev, L.B., Wasserman, N., Wahlested, C. and Zabarovsky, E.R.
TITLE NOT1 flanking sequences: a tool for gene discovery and verification of the human genome
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
MEDLINE 22131767
PUBMED 12136098
REFERENCE 2 (bases 1 to 692)
AUTHORS Zabarovsky, E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden

FEATURES
source
1..692
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NR5-1B14C"

BASE COUNT 61 a 256 c 236 g 129 t 10 others

ORIGIN
Query Match 47.8%; Score 592.2; DB 9; Length 692;
Best Local Similarity 92.9%; Pred. No. 1.7e-69;
Matches 627; Conservative 0; Mismatches 43; Indels 5; Gaps 1;

QY 1 ATGGGACGCCCCCTGGAAACGGCAGCAGCGGCCCGAGGGGGCGCGGAGCGCGCGTGGGCC 60
Db 20 ATGGGACGCCCCCTGGAAACGGCAGCAGCGGCCCGAGGGGGCGCGGAGCGCGCGTGGGCC 79

QY 61 GCAGTCGCGCCTTGGCAGCAGAGCCGCTGCTGCCCTTTTCCCTGGGGGCGCTGTGTGCG 120
Db 80 GCAGTCGCGCCTTGGCAGCAGAGCCGCTGCTGCCCTTTTCCCTGGGGGCGCTGTGTGCG 139

QY 121 GTACCGCTGTGTCCCTGTGCTTTCGTCGTCGGGGTGAGCGGCACAGTGTGTGACCGTG 180
Db 140 GTACCGCTGTGTGCTGTGTGCTGCTGTCGTCGGGGTGAGCGGCACAGTGTGTGACCGTG 199

QY 181 ATGCTGATCGGGCGCTACCGGGACATCGGACACACCAACTTGTACTGGGCGAGCATG 240
Db 200 ATGCTGATCGGGCGCTACCGGGACATCGGACACACCAACTTGTACTGGGCGAGCATG 259

QY 241 GCGGTGTCCGACCTACTCATCTGCTCGGGCTGCGCTTCGACCTGTACCGCTCTGGGCG 300
Db 260 GCGGTGTCCGACCTACTCATCTGCTCGGGCTGCGCTTCGACCTGTACCGCTCTGGGCG 319

QY 301 TCGGGCGCTGTGGGTGTTCGGGCGCTGTCTGTCGGCGCTGTCCCTCTTACGTGGGCGAGGCG 360
Db 320 TCGGGCGCTGTGGGTGTTCGGGCGCTGTCTGTCGGCGCTGTCCCTCTTACGTGGGCGAGGCG 379

Qy 833 GCGGCGCTGCGAGGCGCCCGCCCTCGGGCGG 867
|||
Db 92236 GGGTCCCGACCAAGGTGCGGAAGAAAGGGCGG 92202
|||

RESULT 15

AR168467
LOCUS AR168467 283 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 4 from patent US 6287855.

ACCESSION AR168467

VERSION AR168467.1 GI:17904400

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: US 6287855-A 4 11-SEP-2001;

Location/Qualifiers

1. 283

source

/organism="unknown"

BASE COUNT 27 a 116 c 84 g 56 t

ORIGIN

Query Match 22.8%; Score 283; DB 6; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.5e-28;

Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 TCGGACACACCAACTTGTACCTGGGACGATGGCGGTGTCGACCTACTATCCTGC 265

Db 1 TCGGACACACCAACTTGTACCTGGGACGATGGCGGTGTCGACCTACTATCCTGC 60

Qy 266 TCGGCTGCCGTTCCGACCTGTACCGCCTCTGGCGCTCGCGCCCTGGGTGTCGGGCGGC 325

Db 61 TCGGCTGCCGTTCCGACCTGTACCGCCTCTGGCGCTCGCGCCCTGGGTGTCGGGCGGC 120

Qy 326 TGCTCTGCCGCTGTCCCTCTACCTGGGCGAGGGCTGCACCTAGCCACGCTGCTGCACA 385

Db 121 TGCTCTGCCGCTGTCCCTCTACCTGGGCGAGGGCTGCACCTAGCCACGCTGCTGCACA 180

Qy 386 TGACCGGCTCAGCGTCGAGCGTACCTGGCCATCTCGCGCCGCTCCGCGCCGCGTCT 445

Db 181 TGACCGGCTCAGCGTCGAGCGTACCTGGCCATCTCGCGCCGCTCCGCGCCGCGTCT 240

Qy 446 TGGTCACCGCGCGCGGTCCGCGGCTCATCGCTGTGCTCTG 488

Db 241 TGGTCACCGCGCGCGGTCCGCGGCTCATCGCTGTGCTCTG 283

Search completed: January 1, 2004, 01:09:43

Job time : 4728.39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 13:48:11 ; Search time 380.324 Seconds
(without alignments)
8794.080 Million cell updates/sec

Title: US-09-719-485-2

Perfect score: 1239

Sequence: 1 atggcgagccctggaacgg.....acgtgaagacgatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1239	100.0	1239	21	AAZ45403
2	1239	100.0	1239	22	AAZ45403
3	1239	100.0	1239	22	AAZ45403
4	1239	100.0	1239	24	ABK90132
5	1239	100.0	1239	25	ABZ42842
6	1239	100.0	1506	22	AAI66989
7	1078	87.0	1390	21	AAZ45404
8	915.6	73.9	1203	22	AAZ45448

9	901	72.7	1161	22	AAF83684	Short form of motl
10	901	72.7	3066	21	AAZ45402	Genomic sequence o
11	896.2	72.3	2040	21	AAZ46116	Human G protein co
12	546	44.1	813	22	AAF85447	Nucleotide sequenc
13	501.2	40.5	1179	24	ABQ47146	Oligonucleotide fo
14	501.2	40.5	1179	24	ABQ47147	Oligonucleotide fo
15	437.2	35.3	1179	24	ABQ47148	Oligonucleotide fo
16	437.2	35.3	1179	24	ABQ47149	Oligonucleotide fo
17	283	22.8	283	19	AAV4930	Galanin receptor G
18	283	22.8	283	19	AAV32651	Galanin receptor G
19	283	22.8	283	19	AAV28290	Galanin receptor G
20	283	22.8	283	19	ABK14060	Rat galanin recept
21	264.8	21.4	1092	21	AAZ45405	DNA encoding the p
22	264.8	21.4	1092	22	AAZ45405	Nucleotide sequenc
23	255.8	20.6	1050	21	AAZ61492	cDNA encoding cani
24	234.8	19.0	1063	18	AAZ68662	Pig growth hormone
25	234.8	19.0	1063	18	AAZ68662	Pig growth hormone
26	233.2	18.8	1029	18	AAZ68663	Pig growth hormone
27	233.2	18.8	1029	18	AAZ68663	Pig growth hormone
28	231.6	18.7	1095	21	AAZ45993	Swine growth hormo
29	231.6	18.7	4009	21	AAZ45967	cDNA encoding the
30	229.8	18.5	1092	22	AAZ27800	DNA encoding the m
31	229	18.5	250	25	ACA55761	Rat growth hormone
32	228.8	18.5	870	25	ABZ42674	Pig signalling pat
33	228.8	18.5	1122	18	AAZ68665	Human growth hormo
34	228.4	18.4	3129	18	AAZ68667	Human growth hormo
35	228.4	18.4	3129	18	AAZ68667	Rat growth hormone
36	227.8	18.4	1088	18	AAZ68664	Rat growth hormone
37	227.8	18.4	1088	18	AAZ68664	Human growth hormo
38	227.8	18.4	1101	21	AAA30643	Human G protein-co
39	227.8	18.4	1101	21	AAA30732	DNA encoding human
40	227.8	18.4	1101	21	AAZ51463	Human G protein-co
41	227.8	18.4	1101	21	AAZ51463	Human G protein-co
42	227.8	18.4	1101	22	AAZ30395	Human G-protein co
43	227.8	18.4	1101	25	ACA56550	Human growth-hormo
44	227.2	18.3	1122	18	AAZ69757	Human signalling p
45	226.8	18.3	1092	18	AAZ69760	Human growth hormo
						Rat growth hormone

ALIGNMENTS

RESULT 1
AAZ45403
ID AAZ45403 standard; cDNA; 1239 BP.

XX AAZ45403;

AC AAZ45403;

XX 27-MAR-2000 (first entry)

DE cDNA encoding the motilin receptor splice variant MTL-R1A.

XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-esophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
KW endoscopy; duodenal intubation; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1239

FT FT /*tag= a

FT FT /product= "MTL-R1A"

XX WO9964436-A1.

XX PD 16-DEC-1999.

XX Human; G protein-coupled receptor 38; receptor; GPR38; gene; ds;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;
KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..1239
FT /tag= a
FT /product= "Human G protein-coupled receptor 38 (GPR38)"
XX
PN WO200257791-A2.
XX
XX 25-JUL-2002.
XX
XX 29-NOV-2001; 2001WO-US45219.
PF
XX 29-NOV-2000; 2000US-250251P.
PR
XX 30-NOV-2000; 2000US-250452P.
PR
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
PA
XX Brown JP, Burmer GC, Roush CL, Kulander BG;
PI WPI; 2002-566812/60.
XX P-PSDB; ABG30936.
DR
XX Assay for detecting Alzheimer's disease, Parkinson's disease,
PT ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
PT carcinoma, comprises using a binding partner for G protein coupled
PT receptor 38 -
XX
XX Disclosure; Fig 1; 112pp; English.
PS
CC The present invention relates to a new assay method that involves
CC contacting a binding partner specific for G protein coupled receptor
CC (GPR) 38 with specific cells. The method of the invention is useful for
CC the detection of an increased risk of Alzheimer's disease, Parkinson's
CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
CC glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
CC for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
CC glioblastoma, breast carcinoma, colon carcinoma, lung small cell
CC carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
CC pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
CC to manufacture a medicament able to reduce the symptoms of these
CC diseases. Nucleic acids encoding GPR 38 can also be used to treat the
CC diseases. The present nucleic acid sequence encodes the human G protein-
CC coupled receptor 38 (GPR38) of the invention.
XX
SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;

Query Match 100.0%; Score 1239; DB 24; Length 1239;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAGAGCCCTGGAAACGGACGACGACGCGCCCGAGGGGCGCGAGCGCGCGTGGCCC 60
DB 1 ATGGGAGAGCCCTGGAAACGGACGACGACGCGCCCGAGGGGCGCGAGCGCGCGTGGCCC 60

QY 61 GGGCTGCCCTTGGACAGCGCCGCTGCTCGCCCTTTCCCTTGGGGGCGCTGGTGGCG 120
DB 61 GGGCTGCCCTTGGACAGCGCCGCTGCTCGCCCTTTCCCTTGGGGGCGCTGGTGGCG 120

QY 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
DB 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180

QY 181 ATGCTGATCGGGCGCTACCGGGACATCGGGACACCAACCACTTGTACCTGGGCGAGCATG 240
|||||

RESULT 5
AB242842

Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCCCAATCTCTACAACTCATTTCAAG 1080
 Qy 1081 AAGTACAGAGCGCGCTTTAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGAGCTTC 1140
 Db 1081 AAGTACAGAGCGCGCTTTAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGAGCTTC 1140
 Qy 1141 CACAGAAGCAGGACACTCGGGGGAGTTGCAAGGACACTCGAGGAGACACGGTGGGC 1200
 Db 1141 CACAGAAGCAGGACACTCGGGGGAGTTGCAAGGAGACACTCGAGGAGACACGGTGGGC 1200
 Qy 1201 TACACCGAGACAGCGCTAACGTGAAGACGATGGGATAA 1239
 Db 1201 TACACCGAGACAGCGCTAACGTGAAGACGATGGGATAA 1239

RESULT 6

AAI66989

ID AAI66989 standard; DNA; 1506 BP.

XX AAI66989;

XX 30-JAN-2002 (first entry)

DT Human GPR38 variant GPR38V polypeptide encoding DNA.

DE Human GPR38 variant GPR38V polypeptide encoding DNA.

XX GPR38V; variant; antibacterial; cytostatic; analgesic; antiaesthetic;

KW anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;

KW antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;

KW antiulcer; antiemetic; cardiac; vaccine; ds.

XX Homo sapiens.

OS Key

FT Location/Qualifiers

FT 1..1506

FT /*tag= a

FT /product= "GPR38V"

XX WO200164836-A2.

PN 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US06277.

XX 01-MAR-2000; 2000US-0516315.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Elshourbagy N, Shabon U;

PI WPI; 2001-638956/73.

XX P-PSDB; AAG65822.

XX New human GPR38V polypeptide and polynucleotide, useful for treating

PT e.g. bacterial, fungal, protozoal and viral infections, cancers or

PT allergies, as vaccines, and for identifying agonists and antagonists

PT potentially useful in therapy -

PS Claim 2; Page 26; 32pp; English.

XX This DNA encodes a human GPR38 variant (GPR38V) polypeptide. GPR38V can

CC be expressed by standard recombinant methodology. The polynucleotides and

CC polypeptides are used in the treatment of bacterial, fungal, protozoal

CC and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,

CC diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart

CC failure, hypertension, urinary retentions, osteoporosis, allergies,

CC ulcers, migraine, psychotic and neurological disorders, or dyskinesias.

CC They are also useful for identifying agonists and antagonists that are

CC potentially useful in therapy, as vaccines to induce immunological

CC response in a mammal. The polypeptides may also be used as immunogens to

CC produce antibodies immunospecific for the polypeptides, and to identify

CC membrane bound or soluble receptors.

XX Sequence 1506 BP; 207 A; 534 C; 509 G; 256 T; 0 other;

SQ

Query Match 100.0%; Score 1239; DB 22; Length 1506;
 Best Local Similarity 100.0%; Pred. No. 1.9e-216;
 Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGGGAGAGCCCTGGAAACGGACGAGCGCCCGAGGGGGGGGGAGCGCGCGCTGGCCCC 60
 Db 268 ATGGGAGAGCCCTGGAAACGGACGAGCGCCCGAGGGGGGGGGAGCGCGCGCTGGCCCC 327
 Qy 61 GGGCTGCCCTTTGGACGAGCGCCCTGCTGCGCCCTTTCCCTGGGGGGCGCTGGTGGCG 120
 Db 328 GGGCTGCCCTTTGGACGAGCGCCCTGCTGCGCCCTTTCCCTGGGGGGCGCTGGTGGCG 387
 Qy 121 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 388 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447
 Qy 181 ATGCTGATCGGGGGCTACCGGAGCATGCGGACACCAACCACTTGTACCTGGGACGATG 240
 Db 448 ATGCTGATCGGGGGCTACCGGAGCATGCGGACACCAACCACTTGTACCTGGGACGATG 507
 Qy 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCG 300
 Db 508 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCG 567
 Qy 301 TCGCGCCCTGGGTGTTCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 568 TCGCGCCCTGGGTGTTCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
 Qy 361 TGCACCTACGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Db 628 TGCACCTACGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 687
 Qy 421 TGCAGCCCGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 688 TGCAGCCCGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747
 Qy 481 GTGCTCTGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 748 GTGCTCTGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807
 Qy 541 CAGGACCCCGGACATCTCGGTAGTCCCGGGGCTCAATGGGACCCCGGGGATCGGCTCTCTCG 600
 Db 808 CAGGACCCCGGACATCTCGGTAGTCCCGGGGCTCAATGGGACCCCGGGGATCGGCTCTCTCG 867
 Qy 601 CTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 868 CTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
 Qy 661 GGGCCCGAGACCGCGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 928 GGGCCCGAGACCGCGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987
 Qy 721 CAGTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db 988 CAGTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047
 Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGGGGAGCTGTGGAGACGCGCGGCGCGCG 840
 Db 1048 CTGTGCTCAGCATCTCTACGGCTCATCGGGGGGAGCTGTGGAGACGCGCGGCGCGCG 1107
 Qy 841 CTGCGAGGCGCGCGCGCTCGGGGGGAGAGAGGCGCACCGGAGACCGCTCGCGCGCTCGTG 900
 Db 1108 CTGCGAGGCGCGCGCGCTCGGGGGGAGAGAGGCGCACCGGAGACCGCTCGCGCGCTCGTG 1167
 Qy 901 CTGGTGGTGTCTGCGATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db 1168 CTGGTGGTGTCTGCGATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227
 Qy 961 TACATAAACACGGAAGATTTCGGGATGATGATCTCTCAGTACTTAAACATGCTGCTGCT 1020
 Db 1228 TACATAAACACGGAAGATTTCGGGATGATGATGATCTCTCAGTACTTAAACATGCTGCTGCT 1287

QY 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCCCAATCTCTACACCTCATTTCAAAG 1080
 Db 1288 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCCCAATCTCTACACCTCATTTCAAAG 1347
 QY 1081 AAGTACAGAGCGCGGCTTTAACTCTGCTGCAAGGAAGTCCAGGCGGAGAGGCTTC 1140
 Db 1348 AAGTACAGAGCGCGGCTTTAACTCTGCTGCAAGGAAGTCCAGGCGGAGAGGCTTC 1407
 QY 1141 CACAGAAGCAGGACACTGCGGGGGAAGTTGCAAGGGGACACTGGAGGAGACAGGTGGGC 1200
 Db 1408 CACAGAAGCAGGACACTGCGGGGGAAGTTGCAAGGGGACACTGGAGGAGACAGGTGGGC 1467
 QY 1201 TACACCGAGACAAGCGCTAACTGTAAGACGATGGATAA 1239
 Db 1468 TACACCGAGACAAGCGCTAACTGTAAGACGATGGATAA 1506

RESULT 7
 AA45404
 ID AA45404 standard; cDNA; 1390 BP.
 XX
 AC AA45404;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE cDNA encoding the motilin receptor splice variant MTL-R1B.
 XX
 KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation; ds.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..1161
 FT /*tag= a
 FT /product= "MTL-R1B"
 FT
 XX WO9964436-A1.
 XX
 PD 16-DEC-1999.
 XX
 XX 08-JUN-1999; 99WO-US12773.
 XX
 XX 12-JUN-1998; 98US-0089098.
 PR (MERI) MERCK & CO INC.
 XX
 XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;
 PI
 XX WPI; 2000-105868/09.
 XX
 DR P-PSDB; AAY54146.
 XX
 XX Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions
 XX
 PS Claim 6; Fig 4; 44pp; English.
 XX
 XX The present sequence encodes splice variant MTL-R1B of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility

CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impactation, postoperative ileus, gallstones, infantile colic, irritable
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.
 XX
 SQ Sequence 1390 BP; 190 A; 496 C; 436 G; 268 T; 0 other;
 Query Match 87.0%; Score 1078; DB 21; Length 1390;
 Best Local Similarity 89.1%; Pred. No. 3.6e-187;
 Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
 QY 1 ATGGGAGAGCCCTTGAACGCGACGCGCGCCGAGGGGGCGGAGCGCCGTGGCCCC 60
 Db 1 ATGGGAGAGCCCTTGAACGCGACGCGCGCCGAGGGGGCGGAGCGCCGTGGCCCC 60
 QY 61 GCGCTGCCGCTTGGCGACGAGCGCGCTGCTCGCCCTTTCCCTGGGGGGCGTGGTCCG 120
 Db 61 GCGCTGCCGCTTGGCGACGAGCGCGCTGCTCGCCCTTTCCCTGGGGGGCGTGGTCCG 120
 QY 121 GTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 GTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 181 ATGCTGATCGGGCGCTTACCGGACATGCGGACACACCAACCTTGTACCTTGGGAGCATG 240
 Db 181 ATGCTGATCGGGCGCTTACCGGACATGCGGACACACCAACCTTGTACCTTGGGAGCATG 240
 QY 241 GCGTGTCCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 241 GCGTGTCCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 QY 301 TCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 301 TCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 361 TGCACTTACGCCACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Db 361 TGCACTTACGCCACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 421 TCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 421 TCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 GTGCTCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 GTGCTCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 CAGGACCCCGGCATCTCCGAGTCCCGGGCTCAATGGGACACCGCGGATGCGCTCTCTCG 600
 Db 541 CAGGACCCCGGCATCTCCGAGTCCCGGGCTCAATGGGACACCGCGGATGCGCTCTCTCG 600
 QY 601 CCTCTGCGCTGCTGCGCGCTCTCTGCTGCTGCGGGGGCGCACCGCGCTCCCGCGCTCG 660
 Db 601 CCTCTGCGCTGCTGCGCGCTCTCTGCTGCTGCGGGGGCGCACCGCGCTCCCGCGCTCG 660
 QY 661 GGGCCCGAGACCGCGGAGGCGGGGCTGCTTCAAGCGCGGAAATGCGCGGCGAGCGCCGCG 720
 Db 661 GGGCCCGAGACCGCGGAGGCGGGGCTGCTTCAAGCGCGGAAATGCGCGGCGAGCGCCGCG 720
 QY 721 CAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 CAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGCGGGAGCTTGGAGCAGCGCGGGCGCG 840
 Db 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGCGGGAGCTTGGAGCAGCGCGGGCGCG 840
 QY 841 CTGCGAGGCGCGCGCTCTGCGGGGAGAGAGGCGCACCGCGAGACCGTCCGCGTCTCTG 900

Db 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCCACCGGACACCGTCCGGCTCTG 900
Qy 901 C----- 901
Db 901 CGTAAGTGGAGCGCGCGTGGTTCCAAAGACGCTGCTGCAGTCCGCGCGCGGAGCC 960
Qy 902 ----- 901
Db 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCCAGCTCTGGGGCGCGCTTCCAGCTCCC 1020
Qy 902 -----TGGTGGTGGTTCTGGCATTTATAATTG 929
Db 1021 TTTCCTATTTCGATTCAGCTCCACCGCGGTGGTGGTCTGGCATTTATAATTG 1080
Qy 930 CTGGTTCCTCTCCACGTTGGCAGAACTATTACATAAAACACGGAAGATTGCGGATGAT 989
Db 1081 CTGGTTCCTCTCCACGTTGGCAGAACTATTACATAAAACACGGAAGATTGCGGATGAT 1140
Qy 990 GTACTTCTCAGTACTTTAAACATCGTGGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
Db 1141 GTACTTCTCAGTACTTTAAACATCGTGGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
Qy 1050 CAACCCATCTCTACAACTCTATTCAAGAAGTACAGAGCGCGGCTTTAACTGCT 1109
Db 1201 CAACCCATCTCTACAACTCTATTCAAGAAGTACAGAGCGCGGCTTTAACTGCT 1260
Qy 1110 GCTCGAAGGAAGTCCAGGCGGAGAGGCTTCCACAGAAGCAGGACACTGCGGGGAAGT 1169
Db 1261 GCTCGAAGGAAGTCCAGGCGGAGAGGCTTCCACAGAAGCAGGACACTGCGGGGAAGT 1320
Qy 1170 TGCAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAAAGCGCTAAAGAGAC 1229
Db 1321 TGCAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAAAGCGCTAAAGAGAC 1380
Qy 1230 GATGGGATAA 1239
Db 1381 GATGGGATAA 1390

RESULT 8
AAF85448
ID AAF85448 standard; cDNA; 1203 BP.
AC AAF85448;
XX
XX
DT 23-JUL-2001 (first entry)
XX
XX Nucleotide sequence of a rabbit motilin receptor polypeptide.
DE Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea; 88.
XX
XX Oryctolagus cuniculus.
XX
XX Key Location/Qualifiers
FH 1..1203
FT /*tag= a
FT /note= "motilin receptor"
XX
XX W0200132710-A1.
XX
XX 10-MAY-2001.
XX
XX 25-OCT-2000; 2000WO-US29426.
XX
XX 29-OCT-1999; 99US-0162264.
XX (MERI) MERCK & CO INC.
XX
XX Tan C, McKee K;
XX WPI; 2001-343479/36.
DR

P-PSDB; AAB68477.
Novel polypeptides related to dog and rabbit motilin receptor polypeptide, comprising unique regions from dog and motilin receptor amino acid sequence, useful for identifying compounds for treating diarrhoea in humans -
XX
XX Claim 18; Page 18-19; 42pp; English.
XX
XX The present sequence encodes a rabbit motilin receptor polypeptide. The specification describes a unique sequence present in exon 1 of the dog motilin receptor, which is not present in human or Sphaeroides nophilus 7587 motilin receptor sequences. The unique nucleic acid sequence is useful for measuring the ability of a compound to affect motilin receptor activity. Motilin receptor polynucleotides and polypeptides are used to identify therapeutic compounds which are useful for treating gastrointestinal diseases and disorders such as gastric motility disorders, gastroparesis, irritable bowel syndrome, and diarrhoea.
XX
XX Sequence 1203 BP; 154 A; 423 C; 403 G; 223 T; 0 other;
Query Match 73.9%; Score 915.6; DB 22; Length 1203;
Best Local Similarity 85.3%; Pred. No. 1.2e-157; Indels 36; Gaps 2;
Matches 1058; Conservative 0; Mismatches 144;
Qy 1 ATGGGAGAGCCCTTGGAAACGGCAGCAGCGCCGCGAGGGGCGGAGCGCGCTGGGCC 60
Db 1 ATGGGAGAGCCCTTGGAAACGGCAGCAGCGCCGCGAGGGGCGGAGCGCGCTGGGCC 60
Qy 61 GCCTGCGCGCTTTCGACGAGCGCCGCTGCTCCGCTTTCCTGGGGGCGCTGGTGGCG 120
Db 61 GCCTGCGCGCGCTTTCGACGAGCGCCGCTGCTCCGCTTTCCTGGGACGCTGGTGGCT 120
Qy 121 GTACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGACATGCGGACACACACCACTTGTACCTGGGCGAGCATG 240
Db 181 CTGCTGATCGGGCGCTACCGGACATGCGGACACACCACTTGTACCTGGGCGAGCATG 240
Qy 241 GCGGTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
Db 241 GCGGTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
Qy 301 TCGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 360
Db 301 TCGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 360
Qy 361 TGCACCTAGCCACCGCTGCTGACATGACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 420
Db 361 TGCACCTAGCCACCGCTGCTGACATGACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 420
Qy 421 TGCAGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 480
Db 421 TGCAGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 480
Qy 481 GTGCTGTGGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 540
Db 481 GTGCTGTGGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 540
Qy 541 CAGGAGCGCGCATCTCCGAGTCCCGGCTCAATGGACCGCGGCTGCGGCTGCGGCTGCGGCT 600
Db 541 CAGGAGCGCGCATCTCCGAGTCCCGGCTCAATGGACCGCGGCTGCGGCTGCGGCTGCGGCT 600
Qy 601 CCTCTGCGCTTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 660
Db 601 TCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 627
Qy 661 GGGCGCGAGACCGGAGGCGCGGCTGCTTTCAGCGCGGCTTTCAGCGCGGCTTTCAGCGCG 720
Db 628 GGGCGCG---GAGCGGAGGCGCGGCTTCTGTTTTCAGCGCGGCTTTCAGCGCGGCTTTCAGCGCG 684

QY 721 CAGCTGGGCGGCTGCGTGTATGCTGGGTACACCGGCTACTTCTTCTGCCCTTT 780
 Db |||||
 QY 685 CAGCTGGGCTGTGCGGTATGCTGGGTACACCGGCTACTTCTTCTGCCCTTC 744
 Db |||||
 QY 781 CTGTGCTCAGCATCTCTACGGGCTATCGGGCGGAGCTGTGGAGAGCGCGGCGG 840
 Db |||||
 QY 745 CTCTGCTCAGCATCTCTACGGGCTATCGGGCGGAGCTGTGGCGGGTCTGGGCGCG 804
 QY 841 CTGCGAGCGCGGCGCTGCGGGCGGAGAGGCGCACCGGAGACCGTCCGCTCTG 900
 Db |||||
 QY 805 CTGCGAGCGCGGCGGCGGAGAGGCGCACCGGAGAGGCGCGTCCGCTCTG 864
 QY 901 CTGTGCTGCTGCTGCGCATCTTATATTTGCTGGTTCCTTCCAGTTCGAGAAATCAT 960
 Db |||||
 QY 865 CTGTGCTGCTGCTGCGCATCTTATATTTGCTGGTTCCTTCCAGTTCGAGAAATCAT 924
 QY 961 TACATAACAGGAGATTCGCGGATGATGATCTCTAGTACTTTAATCATCTGCT 1020
 Db |||||
 QY 925 TACATAACAGGAGATTCGCGGATGATGATCTCTAGTACTTTAATCATCTGCT 984
 QY 1021 CTGCAATTTTCTATCTGAGCGCATCTATCAACCAATCTCTACACCTCATTTCAAG 1080
 Db |||||
 QY 985 CTGCAATTTTCTATCTGAGCGCATCTATCAACCAATCTCTACACCTCATTTCAAG 1044
 QY 1081 AAGTACAGAGCGGCGGCTTTTAAACTGCTCTGCTCAAGAGTCCAGCGGAGAGGCTTC 1140
 Db |||||
 QY 1045 AAGTACAGAGCGGCTGCGCGAGACTGCTGCGGAGAGCGGCGGCGGCTGCTG 1104
 QY 1141 CACAGAGAGGAGACACTGCGGGGAAAGTTGACGGGAGACACTGAGAGAGACAGTGGG 1200
 Db |||||
 QY 1105 TCGGGAAGCAGGGGCGCTGAGCAGGAGCTTGCAGGGGAGACACTGCGGAGACACAGCTGGC 1164
 QY 1201 TACACCGAGCAAGCGCTACGTCAAGACGATGGGATA 1238
 Db |||||
 QY 1165 TGCACCGAGACAGCGCTAAACAAAGACGCTGCATA 1202

RESULT 9

AAFP83684
 ID AAFP83684 standard; DNA; 1161 BP.
 AC AAFP83684;
 XX
 XX 23-JUL-2001 (first entry)
 XX
 XX Short form of motilin receptor, GPR-38B isoform encoding DNA.
 XX
 XX zsig33; signal transduction; hormone; enzyme; neural development;
 XX gastric contractility; nutrient uptake; digestive; pancreatic; human;
 XX insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 XX glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R; ds;
 XX G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..1161
 /*tag= a
 /product= "GPR-38B"

WO200138355-A2.

31-MAY-2001.

XX 22-NOV-2000; 2000WO-US932074.

XX 22-NOV-1999; 99US-0166765.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;

DR WPI; 2001-355879/37.
 XX P-PSDB; AAB62653.
 XX
 PT Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 XX secretion, involves contacting a receptor with zsig33 polypeptide -
 PS Disclosure; Page 106-109; 111pp; English.
 XX
 CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
 CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the DNA encoding
 CC the short form of motilin receptor, GPR-38B (one of the two isoforms of
 CC GPR38 which result from alternative splicing). GPR38 has homology to the
 CC human G-protein coupled receptor, GHS-R.
 XX
 SQ Sequence 1161 BP; 126 A; 436 C; 372 G; 227 T; 0 other;

Query Match 72.7%; Score 901; DB 22; Length 1161;
 Best Local Similarity 100.0%; Pred. No. 5.4e-155;
 Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAGCGCCCTGGAAACGCGACGAGCGCGCGGAGGCGCGGAGCGCGCGTGGCCC 60
 Db |||||
 QY 1 ATGGGAGCGCCCTGGAAACGCGACGCGCGCGGAGGCGCGGAGCGCGCGTGGCCC 60
 Db |||||
 QY 61 GCCTGCGCGCTTTCGACGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGTGGTGGCG 120
 Db |||||
 QY 61 GCCTGCGCGCTTTCGACGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGTGGTGGCG 120
 QY 121 GTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db |||||
 QY 121 GTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACACACCACTGTGTACCTGGGCGAGCATG 240
 Db |||||
 QY 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACACCACTGTGTACCTGGGCGAGCATG 240
 QY 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db |||||
 QY 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 QY 301 TCGCGGCGCTTGGGTGTTTCGGGCGCGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTG 360
 Db |||||
 QY 301 TCGCGGCGCTTGGGTGTTTCGGGCGCGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTG 360
 QY 361 TGCACCTAGCCACCGCTGTGACATGACCGCGCTCAGCGCTCAGCGCTTACCTGGCCATC 420
 Db |||||
 QY 361 TGCACCTAGCCACCGCTGTGACATGACCGCGCTCAGCGCTTACCTGGCCATC 420
 QY 421 TGC CGCGCGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db |||||
 QY 421 TGC CGCGCGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 GTGCTCTGGGCGGTGGCGCTGCTCTGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db |||||


```
Db 481 GTGCTCTGGGCGGTGGCGGTCTCTCTGCGGGTCCCTTCTTGTTCCTGGTGGGCGTCGAG 540
Qy 541 CAGACCCCGGCATCTCGTAGTCCCGGGCTCAATGACACCGCGGATCGCTCTCTCG 600
Db 541 CAGACCCCGGCATCTCGTAGTCCCGGGCTCAATGACACCGCGGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCTGTCGCGCTCTCTGCTCTGCGGGGCGCACCGCGCTCCCGCGCTCG 660
Db 601 CCTCTCGCTCTGTCGCGCTCTCTGCTCTGCGGGGCGCACCGCGCTCCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTTCAGCGCGGAATGCCGGCGAGCCCGCGG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTTCAGCGCGGAATGCCGGCGAGCCCGCGG 720
Qy 721 CAGTGGGCGCGCTGCGTGTCTGCTGTGCTGCTACCGCTACTTCTTCTGCTCCCTTT 780
Db 721 CAGTGGGCGCGCTGCGTGTCTGCTGTGCTGCTACCGCTACTTCTTCTGCTCCCTTT 780
Qy 781 CTGTGCTCTCAGCATCTCTACGGCTCATCGGGCGGAGCTGTGGAGCGCGCGCGG 840
Db 781 CTGTGCTCTCAGCATCTCTACGGCTCATCGGGCGGAGCTGTGGAGCGCGCGCGG 840
Qy 841 CTGGAGCGCGCGCTGCGGCGGAGAGCGCACCGCGAGACCGTCCGCTCTG 900
Db 841 CTGGAGCGCGCGCTGCGGCGGAGAGAGCGCACCGCGAGACCGTCCGCTCTG 900
Qy 901 C 901
Db 901 C 901
```

RESULT 10

```
AAZ45402
ID AAZ45402 standard; DNA; 3066 BP.
```

```
AC AAZ45402;
```

```
XX 27-MAR-2000 (first entry)
```

```
XX Genomic sequence of the motilin receptor gene including 5' UTR.
```

```
XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
XX spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
XX functional defect; neurological disorder; scleroderma; colonoscopy;
XX paraneoplastic syndrome; radiation induced dysmotility; diabetes;
XX infection; stress-related motility disorder; psychogenic disorder;
XX gastroparesis; gastro-oesophageal reflux disease; constipation;
XX chronic idiopathic pseudo obstruction; acute faecal impaction;
XX postoperative ileus; gallstones; infantile colic; diarrhoea;
XX irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
XX endoscopy; duodenal intubation; da.
```

```
XX Homo sapiens.
```

```
XX Key Location/Qualifiers
```

```
FT misc_feature 1929..1931
```

```
FT /tag= a
```

```
FT /note= "imperfect donor site"
```

```
FT intron 1930..2728
```

```
FT /tag= b
```

```
FT /note= "intron sequence"
```

```
FT misc_feature 2080..2082
```

```
FT /tag= c
```

```
FT /note= "perfect donor site"
```

```
FT misc_feature 2729..2732
```

```
FT /tag= d
```

```
FT /note= "perfect splice acceptor site"
```

```
XX WO9964436-A1.
```

```
XX 16-DEC-1999.
```

```
XX 08-JUN-1999;
```

```
XX 99WO-US12773.
```

```
XX
```

```
PR 12-JUN-1998; 98US-0089098.
```

```
PA (MERI ) MERCK & CO INC.
```

```
XX
```

```
PI Feigner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
```

```
PI Pong S, Smith RG;
```

```
XX
```

```
DR WPI; 2000-105868/09.
```

```
XX
```

```
PT Novel receptor protein for screening compounds used in treating
```

```
PT irritable bowel syndrome, constipation and other gastric conditions -
```

```
XX
```

```
PS Example 1; Fig 1; 44pp; English.
```

```
XX The present sequence represents the genomic sequence of the motilin
XX receptor gene, including the 5' untranslated region (5' UTR). This gene
XX encodes a G-protein coupled receptor, and is designated MTL-R1 (also
XX GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see AAY54145) and
XX MTL-R1B (see AAY54146). MTL-R1A is a functional seven transmembrane
XX domain form, and MTL-R1B is a truncated five transmembrane domain. The
XX MTL-R1 proteins are used to identify agonists and antagonists which can
XX be used for treating gastric motility disorders, functional defects,
XX disorders secondary to neurological disorders e.g. scleroderma,
XX paraneoplastic syndromes radiation induced dysmotility, diabetes,
XX infections, stress-related motility disorders, psychogenic disorders,
XX gastroparesis, gastro-oesophageal reflux disease, constipation,
XX chronic idiopathic pseudo obstruction, acute faecal impaction,
XX postoperative ileus, gallstones, infantile colic, irritable bowel
XX syndrome, non-ulcer dyspepsia, non-cardiac chest pain and diarrhoea.
XX They can also be used in the preparation for colonoscopy, endoscopy and
XX duodenal intubation.
```

```
XX SQ Sequence 3066 BP; 585 A; 930 C; 926 G; 625 T; 0 other;
```

```
Query Match 72.7%; Score 901; DB 21; Length 3066;
```

```
Best Local Similarity 100.0%; Pred. No. 5.7e-155;
```

```
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGGGAGCGCCCTGGAAACGCGAGCGCGCCCGAGGGGGGGGGGGGGGGGGGGGGGGGG 60
```

```
Db 1029 ATGGGAGCGCCCTGGAAACGCGAGCGCGCCCGAGGGGGGGGGGGGGGGGGGGGGGGGG 1088
```

```
Qy 61 GCGTCGCGCTTGGCGACGAGCGCGCTGTCGCGCCCTTCCCTGGGGGGGGCTGGTGCGG 120
```

```
Db 1089 GCGTCGCGCTTGGCGACGAGCGCGCTGTCGCGCCCTTCCCTGGGGGGGGCTGGTGCGG 1148
```

```
Qy 121 GTACCGCTGTGTGCTGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 180
```

```
Db 1149 GTACCGCTGTGTGCTGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1208
```

```
Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGCATG 240
```

```
Db 1209 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGCATG 1268
```

```
Qy 241 GCGGTGTCCGACCTACTCATCTGCTCGGGCTGCGGCTTCCGACCTGTACCGGCTCTGGCGG 300
```

```
Db 1269 GCGGTGTCCGACCTACTCATCTGCTCGGGCTGCGGCTTCCGACCTGTACCGGCTCTGGCGG 1328
```

```
Qy 301 TCGCGGCGCTTGGGTGTTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
```

```
Db 1329 TCGCGGCGCTTGGGTGTTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1388
```

```
Qy 361 TGCACCTACCGACCTGCTGTCATGACCGCGCTCGAGCGCTACCTGGGCGCATC 420
```

```
Db 1389 TGCACCTACCGACCTGCTGTCATGACCGCGCTCGAGCGCTACCTGGGCGCATC 1448
```

```
Qy 421 TCGCGGCGCTTGGGTGTTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
```

```
Db 1449 TCGCGGCGCTTGGGTGTTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1508
```

```
Qy 481 GTGCTCTGGGCGCTGGCGCTGCTCTCTGCGCGTCCCTTCTTGTTCCTGGTGGGCGTCGAG 540
```

Db 1509 GTGCTCTGGGCGGCTGCTCTCTGCGCGTCCCTTTCTTGTTCCTGGTGGGCGTCGAG 1568
 Qy 541 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 600
 Db 1569 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 1628
 Qy 601 CCTCTGCGCTGCTGCGCGCTCTCTGAGCTCTGCGGGCGCACCGCGCTCCCGCGCTCG 660
 Db 1629 CCTCTGCGCTGCTGCGCGCTCTCTGAGCTCTGCGGGCGCACCGCGCTCCCGCGCTCG 1688
 Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGGATCCCGCGCGAGCGCGCG 720
 Db 1689 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGGATCCCGCGCGAGCGCGCG 1748
 Qy 721 CAGCTGGCGCGCTGCGTGTCTCATCTGTGGGTACACCGCGCTACTTCTTCTGCGCTTT 780
 Db 1749 CAGCTGGCGCGCTGCGTGTCTCATCTGTGGGTACACCGCGCTACTTCTTCTGCGCTTT 1808
 Qy 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCG 840
 Db 1809 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCG 1868
 Qy 841 CTGGAGCGCGCGCGCTCGGGCGGAGAGAGCGCCACCGGAGAGCGTCCGCGTCTG 900
 Db 1869 CTGGAGCGCGCGCGCTCGGGCGGAGAGAGCGCCACCGGAGAGCGTCCGCGTCTG 1928
 Qy 901 C 901
 Db 1929 C 1929

RESULT 11
 ID AAA46116
 AC AAA46116 standard; cDNA; 2040 BP.
 XX
 AC AAA46116;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human G protein coupled receptor hGPR38 (V297K) cDNA SEQ ID NO:129.
 XX
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200022131-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99WO-US24085.
 XX
 PR 13-OCT-1998; 98US-0170496.
 PR 12-NOV-1998; 98US-0108029.
 PR 20-NOV-1998; 98US-0109213.
 PR 27-NOV-1998; 98US-0110060.
 PR 16-FEB-1999; 99US-0120416.
 PR 26-FEB-1999; 99US-0121852.
 PR 12-MAR-1999; 99US-0123944.
 PR 12-MAR-1999; 99US-0123945.
 PR 12-MAR-1999; 99US-0123946.
 PR 12-MAR-1999; 99US-0123948.
 PR 12-MAR-1999; 99US-0123949.
 PR 12-MAR-1999; 99US-0123951.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 30-JUN-1999; 99US-0137567.
 PR 99US-0141448.

PR 27-AUG-1999; 99US-0151114.
 PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 XX Behan DP, Lehmann-Bruinema K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX
 DR WPI; 2000-317986/27.
 DR P-PSDB; AAB02854.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX
 PS Example 2; Page 166-168; 187pp; English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 2040 BP; 350 A; 630 C; 597 G; 463 T; 0 other;
 Query Match 72.3%; Score 896.2; DB 21; Length 2040;
 Best Local Similarity 99.7%; Pred. No. 4.2e-154;
 Matches 898; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 ATGGGAGCGCCCTGGAAACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 Db 1 ATGGGAGCGCCCTGGAAACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 Qy 61 GCCTGCGCGCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 61 GCCTGCGCGCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Qy 121 GTGACCGCTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 180
 Db 121 GTGACCGCTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 180
 Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGCATG 240
 Db 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGCATG 240
 Qy 241 GCGGTGTCGACCTACTCATCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCG 300
 Db 241 GCGGTGTCGACCTACTCATCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCG 300
 Qy 301 TCGCGCGCTGCGGTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCG 360
 Db 301 TCGCGCGCTGCGGTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCG 360
 Qy 361 TGCACCTAGCCACCGCTGTGTCGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
 Db 361 TGCACCTAGCCACCGCTGTGTCGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
 Qy 421 TGCAGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCT 480
 Db 421 TGCAGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCT 480
 Qy 481 GTGCTGTGGCGCTGTCGGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCT 540
 Db 481 GTGCTGTGGCGCTGTCGGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCGCTGTCGGCGCT 540
 Qy 541 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 600

Db 541 CAGGACCCCGGATCTCGTAGTCCCGGCGCTCAATGGACACGGCGGATCGCTCTCTCG 600
 Qy 601 CCTCTCGCTCGTCCCGCTCTCTGCTCTCGGGGCGCACCGCCCTCCCGCGCTCG 660
 Db 601 CCTCTCGCTCGTCCCGCTCTCTGCTCTCGGGGCGCACCGCTCCCGCGCTCG 660
 Qy 661 GGGCCCGAGACCGGAGCGCGGCGCTGTCAGCCGCGAATGCGGCGCGAGCCCGCG 720
 Db 661 GGGCCCGAGACCGGAGCGCGGCGCTGTCAGCCGCGAATGCGGCGCGAGCCCGCG 720
 Qy 721 CAGCTGGGCGGCTCGGCTGTCATGCTGCGGTTCACCGCGCTACTTCTTCTGCCCTTT 780
 Db 721 CAGCTGGGCGGCTCGGCTGTCATGCTGCGGTTCACCGCGCTACTTCTTCTGCCCTTT 780
 Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGCTGTGGAGCAGCGCGGCGG 840
 Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGCTGTGGAGCAGCGCGGCGG 840
 Qy 841 CTGGAGCGCGCGCTCGGGGCGGAGAGAGGCGCACCGGCGAGCGTCCGCTCTG 900
 Db 841 CTGGAGCGCGCGCTCGGGGCGGAGAGAGGCGCACCGGCGAGCGTCCGCTCTG 900
 Qy 901 C 901
 Db 901 C 901

RESULT 12
 AAF85447
 ID AAF85447 standard; DNA; 813 BP.
 XX AC
 XX AAF85447;
 DT 23-JUL-2001 (first entry)
 DE Nucleotide sequence of a dog motilin receptor exon 1.
 XX
 KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea; ss.
 XX
 OS Canis sp.
 XX
 FH Key Location/Qualifiers
 FT 1..813
 CDS /*tag= a
 FT /note= "motilin receptor exon 1"
 FT
 XX WO200132710-A1.
 XX
 PD 10-MAY-2001.
 XX
 PF 25-OCT-2000; 2000WO-US29426.
 XX
 ER 29-OCT-1999; 99US-0162264.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 XX Tan C, McKee K;
 XX WPI; 2001-343479/36.
 DR P-PSDB; AAB68476.
 XX
 PI Novel polypeptides related to dog and rabbit motilin receptor
 XX polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX
 PS Claim 14; Page 17-18; 42pp; English.
 XX
 CC The present sequence represents exon 1 of a dog motilin receptor gene.
 CC The specification describes an unique sequence present in exon 1 of
 CC the motilin receptor, which is not present in human or Sphaeroides
 CC nephus 75E7 motilin receptor sequences. The unique nucleic acid

CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.

XX SQ Sequence 813 BP; 65 A; 316 C; 305 G; 127 T; 0 other;

Query Match 44.1%; Score 546; DB 22; Length 813;
 Best Local Similarity 80.5%; Pred. No. 1.9e-90;
 Matches 705; Conservative 0; Mismatches 105; Indels 66; Gaps 3;

Qy 25 GAGCGCCCGAGGGCGCGGAGCGCGCTGCGCCCGCTGCGCCCTTGGAGAGCGC 84
 Db 4 GGGCGCCCGGAAACAGACGACGCGCGCGAGGGCGCAGCTGCCGTGCGAGAGCGC 63
 Qy 85 CGCTGCTCGGCCCTTCCCTTGGGGGCGCTGGTGGCGGTGACCGCTGTGCTGTGCTG 144
 Db 64 CTGTGCTCGGCCCTTCCCTTGGGGGCGCTGGTGGCGGTGACCGCGCTGTGCTGTGCG 123
 Qy 145 TTGCTCGTGGGGTGAGCGCAACGTGTGACCGTGTGATGTCGGGCGCTACCGGGAC 204
 Db 124 TTGCGGCTGGCGGTGAGCGGCAACCTGTGTGACGGTGTGCTGTGTCGGCGGTACCGGAC 183
 Qy 205 ATGCGGACCAACCACTTGTATCTGGGAGCAATGCGCCGCTGTCGACCTACTCATCTG 264
 Db 184 ATGCGACCAACCAACCTGTATCTGGGAGCAATGCGCCGCTGTCGACCTGCTCATCTG 243
 Qy 265 CTGGGGTGGCGGTTCGACCTGTACCGCTCTGCGGCTCGGGGCGCTGGGTGTTGCGGCG 324
 Db 244 CTGGGGTGGCGCTCGACCTGTACCGCTCTGCGGCTCGGGGCGCTGGGTGTTGCGGCG 303
 Qy 325 CTGCTCTGGCGCTGCTCCCTCTAGTGGGAGGGCTGACCTACGCGACGCTGCTGCAC 384
 Db 304 CTGCTGTGGCGCTGCTCGCTGTACCTGGGAGGGCTGACCTACGCGACGCTGCTGCAC 363
 Qy 385 ATGACCGCGCTCAGCGTTCGAGCGCTACTTGGCCATCTGCGCGCCGCTCCGCGCGCGCTC 444
 Db 364 GTGAGCGCGCTGAGCGTTCGAGCGCTACTTGGCGGTGTCGCGCGCTCCGCGCGCGCG 423
 Qy 445 TTGTCTACCGCGCGCGCTGCGCGCTCATCTGCTGTCTGTGGGCGCTGGCGCTGCTC 504
 Db 424 CTGCTGTCCGGGCGCGCGCTCATCTGCGCGCTCTGCGGCGCTGTCGGCGCTGCTG 483
 Qy 505 TCTGCGGTCCTCTTCTTCTGTTCTGTTGTTCTGTTGTTCTGTTGTTCTGTTGTTCTGTT 564
 Db 484 TCGGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 534
 Qy 565 CCGGCGCTCAATGGCACCGCGCGGATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 624
 Db 535 CCGGCGCTCAACGGCAGCGCGCG----- 557
 Qy 625 TGGCTCTCGGGGCGCGCACCGCGCTCCCGCGCTCGGGGCGCGAGACGCGGAGGCGCGCG 684
 Db 558 -----GCTGGCGCGGGGCGCGCTCCCGCGCGGGGCGCGAG-----GCG 597
 Qy 685 GCGCTGTTCAGCGCGGAAATGCGGCGAGCGCGCGCGAGCTGGGCGGCTGCGGTGTCTATG 744
 Db 598 GCGCTCTTCAGCGCGGAGTCCCGGCGCGCGCGCTGCGAGCTGGGCGGCGCTGCGGTGTCTATG 657
 Qy 745 CTGTGGGTTCACCGCGCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 804
 Db 658 CTCTGGGTTCACCGCGCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 717
 Qy 805 CTCTACCGGGCGGAGCTGTGGAGCAGCGCGCGCGCTTGGAGAGCGCGCGCGCTCGGGG 864
 Db 718 CGCATCGGCGCGAGCTGCGAGGCGCGCGGGGCGCTTGGGGGCGCGCGCGCTCGGGG 777
 Qy 865 CCGGAGAGAGCGCACCGGCGAGCGGTCCGCTCGCTG 900
 Db 778 CGCAGCGGGGCGCACCGCGAGCGCGCTCCGCTCGCTG 813

CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

XX	SQ	Sequence	1179 BP; 139 A; 159 C; 419 G; 462 T; 0 other;
		Query Match	35.3%; Score 437.2; DB 24; Length 1179;
		Best Local Similarity	74.5%; Pred. No. 1.1e-70;
		Matches	550; Conservative 0; Mismatches 188; Indels 0; Gaps 0;
QY	1	ATGGGAGCCCTGGAAAGCGAGCGAGCGGCCCGAGGGGGCGGGAGCGCGCGTGGCCC	60
DB	442	ATGGGTAGTTTTGGAAACGTAGCGATCGTTTCGAGGGGGCGCGGAGTCGTCTGTGGTTTC	501
QY	61	GCCTGCGCGCTTCGACAGAGCGCGCTGCTCGCCCTTCGCCCTGGGGGGCGCTGGTCCGC	120
DB	502	CGTTGTCTTTTCGACAGAGCGTCTGTTCTGTTTTTTTTTTGGGGGGTGGTGTCTG	561
QY	121	GTGACCGCTGTGTGCTGTGCTGTTCTGTCTCGGGGTGAGCGCAACCTGTGTACCTGGGAGCATG	180
DB	562	GTGATCTGTGTGTTGTTGTTCTGTCTGTCTCGGGGTGAGCGGTAACGTGGTGTGATCGTG	621
QY	181	ATGCTGATCGGGCGCTACCGGGACATCGGAGCAACACCAACTGTGTACCTGGGAGCATG	240
DB	622	ATGTTGATCGGGCGTTATCGGGATATCGGATATTATTAATTATTTGGGTAGTAGTATG	681
QY	241	GCCTGTCCGACCTPACTCATCTGCTCGGGTGCCTTCGACCTGTACGCCCTCTGSCGC	300
DB	682	GTCTGTTTCAATTATTTGTTTGGGTGTCTGATTGTTATCGTTTGTGGCGT	741
QY	301	TCGGCGCCCTGGGTGTTTCGGGCGCTCTCTGCGCGCTGTCCCTCTACGTGGGCGAGGCG	360
DB	742	TCGGCGTTTGGGTGTTTCGGGTCTGTTGTTGCTGTTGTTTTTACGTGGGCGAGGCT	801
QY	361	TGCACCTACGGCAGCTGCTGCAATAGACCGCGCTCAGCTCGAGCGCTACCTGGGCCATC	420
DB	802	TGTAATTACGTTACGTTGTGTTATGATATCGCGTTTAGCGTCGAGCGTTATTTGGTTATT	861
QY	421	TGCGCCCGCTCGCGCCCGCGTCTTGCTTACCGGGCGCGGTCGCGCGCTCATCGCT	480
DB	862	TGTCGTTCTGTTTCGCTTCGGTTTGGTTTATTCGGCGTCCGCTTCGGCGTTATCGTT	921
QY	481	GTGCTCTGGGCGGTGGCGTCTCTCTGCGCGTCCCTCTTGTTCCTGTGTGGGCGTCTGAG	540
DB	922	GTGTTTTGGGTCTGTGGCGTTGTTTTTGTTCGTTTTTTTTGTTTTTGGTGGGCGTCGAG	981
QY	541	CAGAACCCCGCATCTCTGATAGTCCCGGGCGCTCAATGGCAACCGCGCGGATCGCGCTCTCG	600
DB	982	TAGGAATTCGGTATTTTTCTAGTTTCGGGTTTTAAATCGTATCGCGCGGATCGTTTTTTCG	1041
QY	601	CTCTCTCGCTCTGCGCGCTCTCTCTGAGTCTCGGGGGCGCACCGCGCTCCCGCGCTCG	660
DB	1042	TTTTTCTGTTCTGCTCTGTTTTTTTGGTTTTTCGGGGCGTTATCGTCTGTTTTCGTCTGCTG	1101
QY	661	GGGCGCCGAGAACCGGAGGCGCGCGCTGTTTCAGCCGCGAATCCCGGCGAGCCCCCGCG	720
DB	1102	GGGTTTCAGATCGGAGGTCGCGGCTGTTTAGTCGCGAATGTCTGCTCGAGTTTCGCG	1161
QY	721	CAGCTGGGCGCTGCGT	738
DB	1162	TAGTTGGGCGTTGCGT	1179

Search completed: December 31, 2003, 22:22:33
Job time : 388.324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:43:42 ; Search time 87.6584 Seconds
(without alignments)
6238.690 Million cell updates/sec

Title: US-09-719-485-2

Perfect score: 1239

Sequence: 1 atgggcagccctggaacgg.....acgtgaagacgatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCUS COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	22.8	283	3	US-08-993-088A-4
2	283	22.8	283	4	US-08-993-424B-4
3	283	22.8	283	4	US-09-603-680-4
4	234.8	19.0	1063	3	US-09-077-675A-1
5	234.8	19.0	1063	4	US-09-077-674-1
6	230	18.6	1029	3	US-09-077-675A-4
7	230	18.6	1029	4	US-09-077-674-4
8	229.8	18.5	1092	3	US-09-077-675A-15
9	229.8	18.5	1092	4	US-09-077-674-15
10	229	18.5	250	4	US-09-016-434-359
11	228.8	18.5	1122	3	US-09-077-675A-9
12	228.8	18.5	1122	4	US-09-077-674-9
13	228.4	18.4	3129	3	US-08-077-675A-14
14	228.4	18.4	3129	4	US-09-077-674-14
15	227.8	18.4	1088	3	US-09-077-675A-6
16	227.8	18.4	1088	4	US-09-077-674-6
17	227.8	18.4	1101	4	US-09-016-434-1148
18	227.8	18.4	1101	4	US-09-170-496D-87
19	227.8	18.4	1101	4	US-09-170-496D-209
20	153	12.3	836	3	US-09-077-675A-11
21	153	12.3	836	4	US-09-077-674-11
22	134.8	10.9	1248	4	US-09-545-944-1
23	132	10.7	1575	3	US-08-858-876A-1
24	132	10.7	1575	3	US-09-472-880-1
25	125.4	10.1	1342	3	US-08-832-399-1
26	125.4	10.1	1342	3	US-09-372-498-1
27	123	9.9	4080	4	US-09-016-434-1346

28	122.8	9.9	1529	3	US-08-958-876A-3	Sequence 3, Appli
29	122.8	9.9	1529	3	US-09-472-880-3	Sequence 3, Appli
30	122	9.8	1285	4	US-09-016-434-1366	Sequence 1366, Ap
31	120.4	9.7	1233	4	US-09-200-090-1	Sequence 1, Appli
32	120	9.7	1535	4	US-09-668-680-12	Sequence 12, Appli
33	119	9.6	1212	4	US-09-170-496D-113	Sequence 113, App
34	119	9.6	1212	4	US-09-170-496D-223	Sequence 223, App
35	115.4	9.3	1176	4	US-09-200-090-3	Sequence 3, Appli
36	109.8	8.9	1164	3	US-08-993-088A-6	Sequence 6, Appli
37	109.8	8.9	1164	4	US-08-993-424B-6	Sequence 6, Appli
38	109.8	8.9	1164	4	US-09-603-680-6	Sequence 6, Appli
39	109.8	8.9	1365	4	US-08-899-112B-27	Sequence 27, Appli
40	108.2	8.7	1219	4	US-08-981-700A-3	Sequence 3, Appli
41	106.8	8.6	1116	3	US-08-993-088A-18	Sequence 18, Appli
42	106.8	8.6	1116	3	US-08-993-088A-19	Sequence 19, Appli
43	106.8	8.6	1116	4	US-08-993-424B-18	Sequence 18, Appli
44	106.8	8.6	1116	4	US-09-603-680-18	Sequence 18, Appli
45	106.8	8.6	1116	4	US-09-603-680-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-08-993-088A-4
; Sequence 4, Application US/08993088A
; Patent No. 6287855
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,088A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...283
; OTHER INFORMATION: cdna probe
US-08-993-088A-4

Query Match 22.8%; Score 283; DB 3; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.8e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGCTGCGACCTACTATCTGTC 265
DB 1 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGCTGCGACCTACTATCTGTC 60

QY 266 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 325
DB 61 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 120

QY 326 TGCTCTCGCGCTGCTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 385
DB 121 TGCTCTCGCGCTGCTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 180

QY 386 TGACCGGCTCAGGCTGAGGCTACCTGGGCATCTCGCGCCGCTCGGCGCGCGTCT 445
DB 181 TGACCGGCTCAGGCTGAGGCTACCTGGGCATCTCGCGCCGCTCGGCGCGCGTCT 240

QY 446 TGGTACCGCGCGCGCTCGCGCGCTCATCGCTGCTCTG 488
DB 241 TGGTACCGCGCGCGCTCGCGCGCTCATCGCTGCTCTG 283

RESULT 2

US-08-993-424B-4
; Sequence 4, Application US/08993424B
; Patent No. 6337206
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Kolakowski, Lee P., Jr.
; TITLE OF INVENTION: MOUSE GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,424B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846NP2
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...283
; OTHER INFORMATION: cdna probe

US-08-993-424B-4
Query Match 22.8%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.8e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGCTGCGACCTACTATCTGTC 265
DB 1 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGCTGCGACCTACTATCTGTC 60

QY 266 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 325
DB 61 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 120

QY 326 TGCTCTCGCGCTGCTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 385
DB 121 TGCTCTCGCGCTGCTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 180

QY 386 TGACCGGCTCAGGCTGAGGCTACCTGGGCATCTCGCGCCGCTCGGCGCGCGTCT 445
DB 181 TGACCGGCTCAGGCTGAGGCTACCTGGGCATCTCGCGCCGCTCGGCGCGCGTCT 240

QY 446 TGGTACCGCGCGCGCTCGCGCGCTCATCGCTGCTCTG 488
DB 241 TGGTACCGCGCGCGCTCGCGCGCTCATCGCTGCTCTG 283

RESULT 3

US-09-603-680-4
; Sequence 4, Application US/09603680
; Patent No. 6544753
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/603,680
; FILING DATE: 26-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; APPLICATION NUMBER: 08/993,088
; FILING DATE: 18-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846 CA
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other

```
/
/
/ NAME/KEY: Other
/ LOCATION: 1...283
/ OTHER INFORMATION: cdna probe
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-603-680-4

Query Match      22.8%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.8e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 TGGGACACACCAACTTGTACCTGGCGAGCATGGCGCTGCGACCTACTCATCTGC 265
Db 1 TGGGACACACCAACTTGTACCTGGCGAGCATGGCGCTGCGACCTACTCATCTGC 60

Qy 266 TCGGCTCGCGTTTCGACCTGTACCGCTCTTGGCGCTCGCGCCCTGGGTGTTGGGCGC 325
Db 61 TCGGCTCGCGTTTCGACCTGTACCGCTCTTGGCGCTCGCGCCCTGGGTGTTGGGCGC 120

Qy 326 TGCTCTCGCGCTGTCCCTCTACCTGGCGAGGCTGCACCTACGCCACGCTGTGCACA 385
Db 121 TGCTCTCGCGCTGTCCCTCTACCTGGCGAGGCTGCACCTACGCCACGCTGTGCACA 180

Qy 386 TGACCGGCTCAGGCTCAGGCTACCTGGCCATCTGCGCCGCTCGCGCCGCTCT 445
Db 181 TGACCGGCTCAGGCTCAGGCTACCTGGCCATCTGCGCCGCTCGCGCCGCTCT 240

Qy 446 TGGTACCGCGCGCGTTCGCGCGCTGCTGCTGCTGCTG 488
Db 241 TGGTACCGCGCGCGTTCGCGCGCTGCTGCTGCTGCTG 283

RESULT 4
US-09-719-485-2
; Sequence 1, Application US/09077675A
; Patent No. 6241299
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-077-675A-1

Query Match      19.0%; Score 234.8; DB 3; Length 1063;
Best Local Similarity 69.3%; Pred. No. 2.3e-36;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 92 CGCCCTTTCCCTGGGGCGCTGGTGGCGGTGACCGCTGTGCTGCTGCTGCTGCTGCTG 151
Db 69 CGCTTTTCCCGACCGCGCTGTGGCGGCGCTCACCGCCACCTTGGCGCTCTTCGTGG 128

Qy 152 TCGGGGTGAGCGGCAACGCTGTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 211
Db 129 TGGGTATCGGGGCAACCTGCTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188

Qy 212 CCACCAAACTTTGTACTGGGCGAGCATGGCGCTGTCCGACCTACTCATCTGCTGGGC 271
Db 189 CCACCAAACTTACTGCTGCGAGTGGCTTCTCCGACCTACTCATCTTCTCTGCA 248

Qy 272 TCGGTTGACCTGTACCGCTCTGGCGCTGCGGCGCTGGGTGCTTGGGCGCGCTGCTCT 331
Db 249 TGGCGCTGACCTTCTCCGCTCTGGCAGTACCGCGCTTGGAACTTGGCAACCTGCTCT 308

Qy 332 GCGCGCTGCTCCCTTACTGCTGGGCGAGGCTGCACCTACGCCACGCTGCTGCACATGACCG 391
Db 309 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTAGCCACAGTGTCTACCATCACCG 368

Qy 392 CGCTCAGCGTCGAGCGCTACTGCGCATCTGCGCGCGCTCGCGCGCGCTCTTGGTCA 451
Db 369 CGCTGAGCGTCGAGCGCTACTTCCGCTGCTGCTGCGCGCGCTGCGCGCGCTGCTCA 428

Qy 452 CCGCGCGCGCGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
Db 429 CCAAGGCGCGGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488

Qy 512 GTCCCTTTCTGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
Db 489 GCGCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 530

RESULT 5
US-09-077-674-1
; Sequence 1, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-077-675A-1
```

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19589P
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-077-674-1

Query Match 19.08; Score 234.8; DB 4; Length 1063;
Best Local Similarity 69.3; Pred. No. 2.3e-38;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 92 CGCCCTTTCCCTGGGGCGCTGTCGCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
DB 69 CGCTTTCCCAACGCCGCTGTGGGGGGTTCACCGCCACTGGTGGCGCTCTTCGTGG 128
QY 152 TCGGGGTGAGCGGCAACGTCGTGACCGTGTGTCGTGTGTCGTGTCGTGTCGTGTCGTG 211
DB 129 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 188
QY 212 CCACCACCACTTGTACTTGGGAGCATGGCGGTGTCGACCTACTCATCTGCTGCGGC 271
DB 189 CCACCACCACTTGTACTTGGGAGCATGGCGGTGTCGACCTACTCATCTGCTGCGGC 248
QY 272 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 331
DB 249 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 308
QY 332 GCGGCTGTCCCTTACGTGGGCGAGGCTGTCACCTACGCAAGCTGTGTCGACATGACCG 391
DB 309 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTACGCAAGCTGTGTCGACATGACCG 368
QY 392 CGCTCAGCGTCGAGCGCTACTTGGCCATCTGCGCGCGCTGCGCGCGCGCTGCTGTGTC 451
DB 369 CGCTCAGCGTCGAGCGCTACTTGGCCATCTGCTTCCGCGCTGCGCGCGCGCTGCTGTGTC 428
QY 452 CCGGCGCGCGCTGTCGCGCGCTCATGCTGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTG 511
DB 429 CCAAGGCGCGGTTAAAGCTGT 488
QY 512 GTCCCTTTCTTTGTTCTTCTGTTGGGCGTTCGAGCAGGACCCCGGCA 553
DB 489 GGCCCATCTTTCGTGCTGTCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 530

RESULT 6
US-09-077-675A-4
Sequence 4, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Fai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/077,675A
APPLICATION NUMBER: 42,452
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-077-675A-4

Query Match 18.68; Score 230; DB 3; Length 1029;
Best Local Similarity 68.6; Pred. No. 2e-37;
Matches 317; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 92 CGCCCTTTCCCTGGGGCGCTGTCGCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
DB 266 CGCTTTCCCAACGCCGCTGTGGGGGGTTCACCGCCACTGGTGGCGCTCTTCGTGG 325
QY 152 TCGGGGTGAGCGGCAACGTCGTGACCGTGTGTCGTGTGTCGTGTCGTGTCGTGTCGTG 211
DB 326 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 385
QY 212 CCACCACCACTTGTACTTGGGAGCATGGCGGTGTCGACCTACTCATCTGCTGCGGC 271
DB 386 CCACCACCACTTGTACTTGGGAGCATGGCGGTGTCGACCTACTCATCTGCTGCGGC 445
QY 272 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 331
DB 446 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 505
QY 332 GCGGCTGTCCCTTACGTGGGCGAGGCTGTCACCTACGCAAGCTGTGTCGACATGACCG 391
DB 506 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTACGCAAGCTGTGTCGACATGACCG 565
QY 392 CGCTCAGCGTCGAGCGCTACTTGGCCATCTGCGCGCGCTGCGCGCGCGCTGCTGTGTC 451
DB 566 CGCTCAGCGTCGAGCGCTACTTGGCCATCTGCTTCCGCGCTGCGCGCGCGCTGCTGTGTC 625
QY 452 CCGGCGCGCGCTGTCGCGCGCTCATGCTGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTG 511
DB 626 CCAAGGCGCGGTTAAAGCTGT 685
QY 512 GTCCCTTTCTTTGTTCTTCTGTTGGGCGTTCGAGCAGGACCCCGGCA 553
DB 686 GGCCCATCTTTCGTGCTGTCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 727

RESULT 7
US-09-077-674-4
Sequence 4, Application US/09077674

	Query Match	18.6%; Score 230; DB 4; Length 1029;	
	Best Local Similarity	68.6%; Pred. No. 2e-37;	
	Matches 317; Conservative 0;	Mismatches 145; Indels 0; Gaps 0;	
Qy	92	CGCCCTTTCCCTGGGGGCGTGTGTCGGGTACCGCTGTGTGCTGTGCTGTCTGTTCTGTCG	151
Db	266	 CGCTCTTTCCCACCACCGGTGTTGGCGGCGTCACCGGCACCTGGCTGGCGCTCTTCGTGG	325
Qy	152	TGGGGTGAGGGCAACGTGTGTACCGTGATGCTGATCGGGCGCTACGGGACATGCCGA	211
Db	326	 TGGGTATCGGGGGCAACTGCTACGATGCTGGTAGTGTCACGTTCCGCGAGATGCCGA	385
Qy	212	CCACCACAACCTGTACTCTGGGACGATGGCCGTCCGACCTACTCATCTGCTCCGGC	271
Db	386	 CCACCACAACCTTACTCTGTCAGCATGGCCTTCTCCGAATACTCATCTCTCTGCA	445
Qy	272	TGCGGTTTCGACCTGTACCGCCTCTGGGGCTCGCGCCCTGGGTGTTGGGGCGCTGCTCT	331
Db	446	 TGCCCCCTGAACTCTTCGGCCTTTGGCAGTACCGGCGTTGGAACTTGGCAACTGCTCT	505
Qy	332	GCCGCCTGTCCCTCTACGTGGGCGAGGGCTGCACCTACGCCACCGTGTGTGCATGACCG	391
Db	506	 GCAAACCTTCCAGTTCTGTAGCCGAGCTGCACCTTAGCCACACAGTGCTCACCATCACCG	565
Qy	392	CGCTCAGCGTCGAGGCTTACTGTGCCATCTGCCGCCGCTCCGGCCCGCGTCTTGGTCA	451

Db 224 CCACACAACTCTACCTGTCCAGCATGGCTTCTCGGATCTGCTCATCTTCTGTGCA 283
 Qy 272 TGCGTTGACACTGTACCGCTCTGGCGCTCGCGCCCTGGGTGTTCGGCGCGCTGCTCT 331
 Db 284 TGCGCTGGACCTGTGCGCTCTGGAGTAGTACCGCCCTGGAATTCGGGACCTGCTCT 343
 Qy 332 GCOCCTGTCCCTCTACGTGGCGAGGCTGACCTACGCCACGCTGCTGCACATGACCG 391
 Db 344 GCAAACTCTCCAGTTGTTCAGCGAGAGCTGACCTAGCCACGGTCTCACCATCACC 403
 Qy 392 CGCTAGGCTGAGCGCTACTGCGCATCTGCCGCCCGCTCGCGCCCGGCTCTGTGTGCA 451
 Db 404 CGCTAGGCTGAGCGCTACTGCGCATCTGCCGCCCGCTCGCGCCCGGCTCTGTGTGCA 463
 Qy 452 CCGGCGCGGCTCGCGCGCTCATGCTGTGCTCTGGCGCGGCTGCTCTCTCGCG 511
 Db 464 CTAGGGCGCGGCTGAACTGTGTCTCTGTGCTCTGGCGCGGCTGCTCTCTCGCG 523
 Qy 512 GTCCCTTCTTGTCTCTGGTGGGCTCGAGCAGGACCGCGGCTCTCGGTAGTCCCGGCG 571
 Db 524 GGCCCATCTTCGTGCTGTGGCGTGGAGCAGCAAAACGGCA----- 565
 Qy 572 TCAATGCGACCGCGCGANTGCTCTCTCGCTCTCGCTCTCGTCTCGCGCGCTCTCTGTGCTCT 631
 Db 566 ----- 572
 Qy 632 CGCGGGCGCACCGCTCTCGCGCTCGCGCGCGGCGCGAGCGCGGAGCGCGCGCTGT 691
 Db 573 CGGGACACCAACAAATGCGCGCGC-----ACCGAGTTGCGTGTGCGTCTGGGCTG 625
 Qy 692 TCAGCCCGGAATCGCGCGCGAGCGCGCGAGCTGGCGGCTGCTGTGCTGTGCTGTGG 751
 Db 626 TCACC----- 643
 Qy 752 TCACACCGCTACTTCTCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 811
 Db 644 TGTCAGCGCTCTTCTCTTCTACCGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 703
 Qy 812 GCGGGAGCTGTGAGCAGCGCGCGCGCTGCGAGCGCGCGCTCGCGCGCGGAGGA 871
 Db 704 GGAGGAAGCTATGCGGAGACGCGG-----AGATGAGCGGTGGCGCGCTCTCGCGGACC 760
 Qy 872 GAGGCCACCGCGACCGCTCGCGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 931
 Db 761 AGAACACACAGCAGCAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 820
 Qy 932 GGTGCGCTTCCAGTTGGGAGATATTTACATAAACAAGAGATTCGCGGATGATGT 991
 Db 821 GGCTGCGCTTCCAGTTGGGAGATATTTACATAAACAAGATTCGCGGATGATGT 880
 Qy 992 ACTTCTCT-----CAGTACTTTAAACATCGCTCTGCACTTTTCTATCTGAGCG 1042
 Db 881 AGATCGCTCAGATCAGCAGTGTGCACTTGAACCTGTGTCTTGTCTCTTCTACTCAGCG 940
 Qy 1043 CATCTATCAACCCCAATCTCTACAACTCTATTTCAAGAAAGTACAGACCGCGCGCTTTA 1102
 Db 941 CTGCACTCAACCCCAATCTCTACAACTCTATTTCAAGAAAGTACAGACCGCGCGCTTTA 1000
 Qy 1103 AACTGCT 1109
 Db 1001 AACTGCT 1007

RESULT 9
 US-09-077-674-15
 ; Sequence 15, Application US/09077674
 ; Patent No. 6531314
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feighner, Scott D.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Liberator, Paul A.

APPLICANT: Schaeffer, James M.
 APPLICANT: Van Der Ploeg, Leonardus
 TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merck & Co., Inc.
 STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 CITY: Rahway
 STATE: NJ
 COUNTRY: USA
 ZIP: 07065-0900
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/077,674
 FILING DATE: 3-JUN-1998
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Cocuzzo, Anna L.
 REGISTRATION NUMBER: 42,452
 REFERENCE/DOCKET NUMBER: 19589P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 732-594-1273
 TELEFAX: 732-594-4720
 TELEX:
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1092 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-077-674-15

Query Match 18.5%; Score 229.8; DB 4; Length 1092;
 Best Local Similarity 57.7%; Pred. No. 2.2e-37;
 Matches 593; Conservative 0; Mismatches 302; Indels 132; Gaps 5;
 Qy 92 CGCCCTTCCCTGGGCGCTGTCGCGGTGACCGCTGTGTCGCTGTGTCGCTGTGTCG 151
 Db 104 CGCTGTTCCTCCGCTCCGCTGCTGGCAGCGCTACCGCCACCTCGCTGGCGCTCTT 163
 Qy 152 TCGGGGTGAGCGCAACGCTGATGCTGATCGGGCGCTACCGGGACATCGGA 211
 Db 164 TGGGCACTCAGGCAACCTGCTCACTATGCTGGTGTGCTTCCGCGAGCTGGCA 223
 Qy 212 CACACCAACCTGTGTACTGTGGGAGCATGGCGGTGTCCGACCTTCTCTCTGCTGGGC 271
 Db 224 CCACCAACCACTTACTCTGTCAGCATGGCTTCTCGGATCTGCTCATCTTCTGTGCA 283
 Qy 272 TGCCTGTGACCTGTACCGCTCTGGGCTCGCGCGCTGGGTGTGGGGCGCTGCTCT 331
 Db 284 TGCCTGTGACCTGTGCTGGCGCTTGGCAGTACCGCGCTTGGAACTTGGCGACCTGCTCT 343
 Qy 332 GCGCGCTGCTCTTACTGTGGCGAGGCTGACACCTACGCCACGCTGCTGCACATGACCG 391
 Db 344 GCAAACTCTTCCAGTTGTTCAGCGAGAGCTGACACTTACGCCAGGCTCTCACCATCACC 403
 Qy 392 CGCTAGGCTGAGCGCTACTGTCGCGCATCTGCGCGCGCTTCCGCGCGCGCTTGTGTGCA 451
 Db 404 CGCTAGGCTGAGCGCTACTTTCGCGCATCTGCTTCTCTTCTGCGGGCAAGGTGGTGA 463
 Qy 452 CCGGCGCGCGCTCGCGCGCTCATGCTGTGCTCTGCGCGCGCTGGCGCTGCTCTCTGCGCG 511
 Db 464 CTAAGGGCGCGGTGAAGCTGGTGTCTCTTGTCTATCTGGGCGCGGTGCTTCTGACGCGCG 523
 Qy 512 GTCCCTTCTTGTCTGTGGGCGTTCGAGCAGGACCGCGCGCATCTCCGTAGTCCCGGCG 571

```

Db 524 GGGCCATCTTCGTGCTGGTGGCGTGAGACCAAAACGGCA----- 565
Qy 572 TCAATGGCACCGCGCGATCGCTCTCGCTCGCTCGCGCGCTCTCTGCGTCT 631
Db 566 -----CAGATCC 572
Qy 632 CGCGGGGCGCACCGCGTCCCGCGTGGGGGCGGAGACCGCGGCGCGTGT 691
Db 573 CGGGACACCAACGAATGCGCGCC-----ACCGAGTTGCGTCTGCGCTG 625
Qy 692 TCAGCGCGGAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 751
Db 626 TCACC-----GTCAATGGTGTGG 643
Qy 752 TCACCACCGCTACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 811
Db 644 TGTCCAGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
Qy 812 GCGCGGAGCTGTGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGA 871
Db 704 GAGGAGAGCTATGCGCGGAGACGCGG---AGATGCGCGCGTGGCGCGCT 760
Qy 872 GAGGCGCACCGGAGACGCTCGCGTCTCTCTCTCTCTCTCTCTCTCTCT 931
Db 761 AGAACACACAGCAGCAGTGAAGATGCTCTCTCTCTCTCTCTCTCTCTCT 820
Qy 932 GGTGTGCTCTTCCAGCTTGGCAGATCAATTAACACGGAAGATTCCCGGATGAT 991
Db 821 GGTGTGCTCTTCCAGCTTGGGAGATACCTCTTTTCCAGTCTCTCGAGCGTCTCT 880
Qy 992 ACTTCTCT-----CAGTACTTTAACTCGTCTCTCTCTCTCTCTCTCT 1042
Db 881 AGATCGCTCAGATCAGCAGCAGTACTGCAACCTGTGTCTCTCTCTCTCT 940
Qy 1043 CATCTATCAACCACTCTTACACCTCAATTTCAAGAAGTACAGCGCGCGCTTTA 1102
Db 941 CTGCCATCAACCCCACTCTCTACACATCATGTCCAAGAAGTACCGGGTGG 1000
Qy 1103 AACTGCT 1109
Db 1001 AACTGCT 1007

```

```

RESULT 10
US-09-016-434-359
; Sequence 359, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

```

```

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 359:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: TBP1NOT01
; CLONE: 2018536
; US-09-016-434-359

Query Match 18.5%; Score 229; DB 4; Length 250;
Best Local Similarity 98.8%; Pred. No. 2.6e-37;
Matches 240; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 997 TCTCACTACTTTAACTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1056
Db 1 TCTCACTACTTTAACTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Qy 1057 ATCTCTCTCAACCTCTATTTCAAAGAAGTACAGAGCGCGCGCTTTAACT 1116
Db 61 ATCTCTCTCAACCTCTATTTCAAAGAAGTACAGAGCGCGCGCTTTAACT 120
Qy 1117 AGGAAGTCCAGGCGGAGAGGCTTTCCACAGAAGCAGGAGACACTCGGGGG 1176
Db 121 AGGAGTCCAGGCGGAGAGG-TTCCACAGAAGCAGGAGACACTCGGGGG 179
Qy 1177 GACACTGGAGGAGACACGGTGGGCTTACCGGAGCAAGCGCTTAACGTGA 1236
Db 180 GACACTGGAGGAGACACGGTGGGCTTACCGGAGCAAGCGCTTAACGTGA 239
Qy 1237 TAA 1239
Db 240 TAA 242

```

```

RESULT 11
US-09-077-675A-9
; Sequence 9, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```


Search completed: January 1, 2004, 02:55:55
Job time : 92.6584 secs

THIS PAGE BLANK (USPTO)

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1239	100.0	1239	13	US-10-206-677-1	Sequence 1, Appl	
2	1239	100.0	1239	15	US-10-225-567A-472	Sequence 472, App	
3	1239	100.0	1239	15	US-10-230-078-13	Sequence 13, Appl	
4	1239	100.0	1239	15	US-10-230-078-14	Sequence 14, Appl	
5	896.2	72.3	2040	12	US-10-417-820A-129	Sequence 129, App	
6	891.4	71.9	2040	12	US-10-417-820A-151	Sequence 151, App	
7	234.8	19.0	1063	13	US-10-303-204A-1	Sequence 1, Appl	
8	230	18.6	1029	13	US-10-303-204A-4	Sequence 4, Appl	
9	229.8	18.5	1092	13	US-10-303-204A-15	Sequence 15, Appl	
10	228.8	18.5	870	15	US-10-225-567A-139	Sequence 139, App	
11	228.8	18.5	1122	13	US-10-303-204A-9	Sequence 9, Appl	
12	228.4	18.4	3129	13	US-10-303-204A-14	Sequence 14, Appl	
13	227.8	18.4	1088	13	US-10-303-204A-6	Sequence 6, Appl	
14	227.8	18.4	1101	13	US-10-276-392-22	Sequence 22, Appl	
15	227.8	18.4	1101	15	US-10-251-385-87	Sequence 87, Appl	


```

601 CCTCTGCCCTTGTGCGCGCTCTCTCTGCTCTCGGGGCGGCGACGCGCTCCCCGCGCTCG 660
661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTGAGCCGCGAAATGCCCGCGGAGCGCGCG 720
661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTGAGCCGCGAAATGCCCGCGGAGCGCGCG 720
721 CAGCTGGGCGCGCTGCGTGTGATCTGCTGTGGGTGACCAACCGCTACTTCTTCTGCGCTTT 780
721 CAGCTGGGCGCGCTGCGTGTGATCTGCTGTGGGTGACCAACCGCTACTTCTTCTGCGCTTT 780
781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCG 840
781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCG 840
841 CTGGAGCGCGCGCTCTCGGGGCGGAGAGGCGCACCGGAGAGCGCTCCGCGTCTG 900
841 CTGGAGCGCGCGCTCTCGGGGCGGAGAGGCGCACCGGAGAGCGCTCCGCGTCTG 900
901 CTGGTGTGTTCTTCTGAGCGCATTTATAATTTGCTGTGGTTCCTTCCAGCTTCAAGATCAT 960
901 CTGGTGTGTTCTTCTGAGCGCATTTATAATTTGCTGTGGTTCCTTCCAGCTTCAAGATCAT 960
961 TACATAACACGAGGATTCGCGGATGATGATCTCTCTCAGTACTTTTAACTCTGCT 1020
961 TACATAACACGAGGATTCGCGGATGATGATCTCTCTCAGTACTTTTAACTCTGCT 1020
1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACTCATTTCAAAG 1080
1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACTCATTTCAAAG 1080
1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCAAGGAAGTCCAGGCGGAGAGGCTTC 1140
1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCAAGGAAGTCCAGGCGGAGAGGCTTC 1140
1141 CACAGAGCAGGACACTCGGGGGAAGTTGCAAGGAGACACTGGAGGAGACACGCTGGC 1200
1141 CACAGAGCAGGACACTCGGGGGAAGTTGCAAGGAGACACTGGAGGAGACACGCTGGC 1200
1201 TACACCGAGACAAGCGCTTAAGTCAAGACGATGGATAA 1239
1201 TACACCGAGACAAGCGCTTAAGTCAAGACGATGGATAA 1239

RESULT 3
US-10-290-078-13
; Sequence 13, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17492 or 58874
; FILE REFERENCE: MF12001-288P1(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-290-078-13

Query Match 100.0%; Score 1239; DB 15; Length 1239;
Best Local Similarity 100.0%; Pred. No. 2.5e-293;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAGCGCCCTGGAACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 1 ATGGGAGCGCCCTGGAACGCGAGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 61 GCGCTGCGCGCTTGCAGCAGCGCGCTGCTGCGCTTTTCCCTTGGGGGCGCTGGTGGCG 120

```

```

61 GCGCTGCGCGCTTGCAGCAGCGCGCTGCTGCGCTTTTCCCTTGGGGGCGCTGGTGGCG 120
121 GTACACCGCTGTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
121 GTACACCGCTGTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
181 ATGCTGATCGGGGCGCTACCGGGGACATGCGGACCAACCAACCTTGTACCTTGGGAGCATG 240
181 ATGCTGATCGGGGCGCTACCGGGGACATGCGGACCAACCAACCTTGTACCTTGGGAGCATG 240
241 GCGCTGTGCGACCTACTCATCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
241 GCGCTGTGCGACCTACTCATCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
301 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360
301 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360
361 TGCACCTACGCCACCGCTGTGCAATGACCGCGCTCAGCGTTCAGCGCTACCTGGCCATC 420
361 TGCACCTACGCCACCGCTGTGCAATGACCGCGCTCAGCGTTCAGCGCTACCTGGCCATC 420
421 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 480
421 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 480
481 GTGCTCTGGGCGCTGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
481 GTGCTCTGGGCGCTGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
541 CAGGACCGCGCATCTCCGAGTCCCGGCGCTCAATGGCACCGCGCGCATCGCTCCCTCG 600
541 CAGGACCGCGCATCTCCGAGTCCCGGCGCTCAATGGCACCGCGCGCATCGCTCCCTCG 600
601 CTTCTGCGCTGCTGCGCGCTGCTGCTGCGGCGCGCACCGCGCTGCTGCGCGCTGCTG 660
601 CTTCTGCGCTGCTGCGCGCTGCTGCTGCGGCGCGCACCGCGCTGCTGCGCGCTGCTG 660
661 GGGCCCGAGACCGCGGAGCGCGCGCTGCTGCTGCGGCGCGCACCGCGCTGCTGCGCG 720
661 GGGCCCGAGACCGCGGAGCGCGCGCTGCTGCTGCGGCGCGCACCGCGCTGCTGCGCG 720
721 CAGCTGGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
721 CAGCTGGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCGCG 840
781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCGCG 840
841 CTGGAGCGCGCGCTCTCGGGGCGGAGAGGCGCACCGGAGAGCGCTCCGCGTCTG 900
841 CTGGAGCGCGCGCTCTCGGGGCGGAGAGGCGCACCGGAGAGCGCTCCGCGTCTG 900
901 CTGGTGTGTTCTTGGCATTTATAATTTGCTGTGGTTCCTTCCAGCTTGGCAGATCAT 960
901 CTGGTGTGTTCTTGGCATTTATAATTTGCTGTGGTTCCTTCCAGCTTGGCAGATCAT 960
961 TACATAACACGGAAGATTCGCGGATGATGATCTCTCTCAGTACTTTTAACTCTTAACT 1020
961 TACATAACACGGAAGATTCGCGGATGATGATCTCTCTCAGTACTTTTAACTCTTAACT 1020
1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACTCATTTCAAAG 1080
1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACTCATTTCAAAG 1080
1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCAAGGAAGTCCAGGCGGAGAGGCTTC 1140
1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCAAGGAAGTCCAGGCGGAGAGGCTTC 1140
1141 CACAGAGCAGGACACTCGGGGGAAGTTGCAAGGAGACACTGGAGGAGACACGCTGGC 1200

```


Qy 181 ATGCTGATCGGGCGCTACCGGACATCGGACCATCGGACCACTTGTACCTGGGAGCATG 240
 Db 181 ATGCTGATCGGGCGCTACCGGACATCGGACCATCGGACCACTTGTACCTGGGAGCATG 240
 Qy 241 CCGGTGTCGACCTACTCATCTGCTCGGGTTCGGGTTCGACCTGTACCGCTCTGGGCG 300
 Db 241 CCGGTGTCGACCTACTCATATGCTCGGGTTCGGGTTCGACCTGTACCGCTCTGGGCG 300
 Qy 301 TCGGGCGCTCGGTGTTTCGGGCGCTGCTGTCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 301 TCGGGCGCTCGGTGTTTCGGGCGCTGCTGTCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Qy 361 TGCACCTACGACCGCTGCTGCACATGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Db 361 TGCACCTACGACCGCTGCTGCACATGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Qy 421 TCGCGCGCTGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 421 TCGCGCGCTGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Qy 481 GTGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 GTGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Qy 541 CAGGACCCCGGCTATCCGTAGTCCCGGCGCTCAATGCGACCGCGCGATCGCTCTCTCG 600
 Db 541 CAGGACCCCGGCTATCCGTAGTCCCGGCGCTCAATGCGACCGCGCGATCGCTCTCTCG 600
 Qy 601 CCTCTGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 601 CCTCTGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Qy 661 GGGCCGAGACCGCGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 661 GGGCCGAGACCGCGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Qy 721 CAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db 721 CAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCTCATCGGGCTCATCGGGCTC 840
 Db 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCTCATCGGGCTCATCGGGCTC 840
 Qy 841 CTGGAGCGCGCGCTGCTGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAG 900
 Db 841 CTGGAGCGCGCGCTGCTGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAG 900
 Qy 901 CTGGTGGTG 909
 Db 901 CGTAAGTTG 909

RESULT 7
 US-10-303-204A-1
 ; Sequence 1, Application US/10303204A
 ; Publication No. US20030166144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feigner, Scott D.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Liberator, Paul A.
 ; APPLICANT: Schaeffer, James M.
 ; APPLICANT: Van Der Ploeg, Leonardus H. T.
 ; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
 ; FILE REFERENCE: 19589PCA
 ; CURRENT APPLICATION NUMBER: US/10/303,204A
 ; PRIOR FILING DATE: 2002-11-25
 ; PRIOR APPLICATION NUMBER: 09/077,674
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: PCT/US96/19445

; PRIOR FILING DATE: 1996-12-10
 ; PRIOR APPLICATION NUMBER: 60/018,962
 ; PRIOR FILING DATE: 1996-06-06
 ; PRIOR APPLICATION NUMBER: 60/008,582
 ; PRIOR FILING DATE: 1995-12-13
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1063
 ; TYPE: DNA
 ; ORGANISM: sus scrofa
 ; US-10-303-204A-1
 Query Match 19.0%; Score 234.8; DB 13; Length 1063;
 Best Local Similarity 69.3%; Pred No. 8.8e-48;
 Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
 Qy 92 CGCCCTTTCCCTCGGGGCGCTGTCGGGTGACCGGTGATCGGGCGCTACCGGACATCGGCA 211
 Db 69 CGCTCTTCCCACTTTCGGGCGCTGTCGGGCGCTACCGGACCTGCTGCTGCTGCTGCTGCTG 128
 Qy 152 TCGGGGTGAGCGGCAACGTGTCACCGGTGATCGGGCGCTACCGGACATCGGCA 211
 Db 129 TGGGTATCGGGGCAACCTGCTCACGATGCTGATGTCACGCTTCCCGGAGATCGCA 188
 Qy 212 CCACCACTTGTACCTGGGCGATGCGCGTGTCCGACCTACTCATCTGCTCGGC 271
 Db 189 CCACCACTTGTACCTGGGCGATGCGCGTGTCCGACCTACTCATCTGCTCGCA 248
 Qy 272 TCGGCTTGCACCTGTACCGCTCTGCGGCTGCGGCGCTGCGGCTGCGGCTGCTGCTCT 331
 Db 249 TCGGCTTGCACCTTTCGGGCTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTCT 308
 Qy 332 GCGGCTTGCCTTCTAGTGGGCGAGGCTGACCTACGCACTGCTGCTGCTGCTGCTGCTGCTG 391
 Db 309 GCAAACTTTCCTGTTAGCGAGAGCTGCACTACGCACTGCTGCTGCTGCTGCTGCTGCTG 368
 Qy 392 CGCTAGCGTCGAGCGCTACCTGCGCATCTGCGGCGCTGCGGCGCTGCGGCGCTGCTGCTGCTG 451
 Db 369 CGCTAGCGTCGAGCGCTACCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 428
 Qy 452 CCGGCGCGCGCTGCGGCGCTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
 Db 429 CCAAGCGCGGTTAAAGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
 Qy 512 GTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
 Db 489 GGCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 530

RESULT 8
 US-10-303-204A-4
 ; Sequence 4, Application US/10303204A
 ; Publication No. US20030166144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feigner, Scott D.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Liberator, Paul A.
 ; APPLICANT: Schaeffer, James M.
 ; APPLICANT: Van Der Ploeg, Leonardus H. T.
 ; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
 ; FILE REFERENCE: 19589PCA
 ; CURRENT APPLICATION NUMBER: US/10/303,204A
 ; PRIOR FILING DATE: 2002-11-25
 ; PRIOR APPLICATION NUMBER: 09/077,674
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: PCT/US96/19445
 ; PRIOR FILING DATE: 1996-12-10
 ; PRIOR APPLICATION NUMBER: 60/018,962
 ; PRIOR FILING DATE: 1996-06-06

[illegible]

Qy		513 TCCCTTCTTGTTCTCGGTGGGCGTCGAGCAGGACCCCGGCATC 555
Db		528 GCCCATCTTCGTGCTAGTCGGGTGGAGCACGAGAAGGCCACC 570

Search completed: January 1, 2004, 06:44:54
Job time : 617.08 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:38:58 ; Search time 2904.98 Seconds
(without alignments)
10366.072 Million cell updates/sec

Title: us-09-719-485-2
Perfect score: 1239
Sequence: 1 atgggagcccttggaagg.....acgtgaagcagatgggataa 1239

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estopl: *
7: em_estro: *
8: em_estr: *
9: gb_estl: *
10: gb_estc: *
11: gb_estc: *
12: gb_estc: *
13: gb_estc: *
14: gb_estc: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_man: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gssl: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	230	18.6	4435	11 AK049671	AK049671 Mus muscu
2	226.8	18.3	500	10 BF603623	BF603623 269181 MA
3	219.2	17.7	608	14 BY724644	BY724644 BY724644
4	186.4	15.0	843	13 BU553576	BU553576 AGENCOURT

C	5	173.4	14.0	849	13	BU568940
C	7	141.6	11.4	504	28	AZ451922
	8	127	10.3	855	12	BI757121
	9	125.4	10.1	1010	12	BM925480
	10	125.4	10.1	634	10	BG714306
	11	125.2	10.1	1055	12	BM808829
	12	124.8	10.1	540	12	BM737726
	13	124	10.0	1037	12	BM543497
	14	122.8	9.9	1144	12	BM546496
C	15	119	9.6	941	29	CNS036WH
	16	117.8	9.5	1603	12	BM808823
	17	111.8	9.0	824	9	AU244232
	18	111.4	9.0	738	12	BI600170
	19	111.4	9.0	426	14	CB695276
	20	111.2	9.0	574	14	CB608597
	21	109.6	8.8	928	12	BI758259
	22	109.2	8.8	993	13	BO721315
	23	106.4	8.6	448	10	BF323227
	24	105.6	8.5	3277	11	AK036756
	25	105	8.5	2760	11	AK081073
	26	104.8	8.5	427	13	BY284735
	27	104.6	8.4	828	12	BI597845
	28	103.6	8.4	660	14	BY728041
	29	102.6	8.3	866	14	CD246184
	30	102.4	8.3	980	12	BM543468
	31	102.2	8.2	419	13	BY273314
	32	102.2	8.2	994	14	BY705540
	33	99	8.2	1189	11	AK005368
	34	98.8	8.0	397	14	CB772331
	35	98.4	7.9	1918	11	AK053776
C	36	98.4	7.9	751	14	CB154463
	37	97	7.8	1233	14	CA975828
	38	97	7.8	785	12	BI754749
C	39	96.8	7.8	843	29	CNS048GJ
	40	96.8	7.8	670	29	AG046172
	41	96.8	7.8	720	12	BI753905
	42	96.6	7.8	1307	11	CNSLTIIBD
	43	96.6	7.8	555	10	BE751626
	44	95	7.7	825	9	AU079556
	45	95	7.7	463	13	BY257194
				502	13	BX280512

ALIGNMENTS

RESULT 1	AK049671	4435 bp	linear	HTC 05-DEC-2002
LOCUS	Mus musculus	12 days embryo spinal cord cDNA	RIKEN full-length	
DEFINITION	enriched library, clone: C530020122	product: GROWTH HORMONE		
	SECRETAGUE RECEPTOR TYPE 1 homolog [Rattus norvegicus], full			
	Insert sequence.			
ACCESSION	AK049671			
VERSION	AK049671.1	GI:26340405		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K.,			
TITLE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
JOURNAL	High-efficiency full-length cDNA cloning			
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)			
PUBMED	99279253			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K.,			
TITLE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
JOURNAL	Normalisation and subtraction of cap-trapper-selected cDNAs to			
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes			
	Genome Res. 10 (10), 1617-1630 (2000)			
	20499374			

PUBMED REFERENCE AUTHORS	11042159 3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4	FEATURES SOURCE	Location/Qualifiers 1. .4435 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:C530020122" /db_xref="taxon:10090" /clone="C530020122" /tissue_type="spinal cord" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="12 days embryo" 252. .1346 /notes="unnamed protein product; GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 homolog [Rattus norvegicus] (SWISSPROT O08725, evidence: FASTV, 99.5%ID, 100%length, match=1092) putative" /codon_start=1 /protein_id="BAC33866.1" /db_xref="GI:26340406" /translation="MNATPSEPEPNVTLDLDWDASPGNDSLSDELLPLFPAPLLAGVTATCVALLPVGIGSNLLMLVVSRELRITTNLYLSSWAFSDLIIFLCPDLVRLWQPRNFGDLICKLQFVSECTATVITITLVSVERFAICFPRAKVVVTKGRVKLVILVIAVAFCSAGPIFVLVGEHENGDPDTRNECRATEFVRGLLTVMMVSSVFFFLPVCTVLYSLIGLRWRGDAVGSRLDRQKQTVKMLAVVVFAPILCWLPFHVGRLYFSKSPSPGSLTAQISOYCNLVSFVLFSAAINPILYIMSKRYVAVFKLLGPESFSQRKSLTKDESSRAWTSSINT"
TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001) 21085660 11217851 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 4435) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.	BASE COUNT ORIGIN Query Match 18.6%; Score 230; DB 11; Length 4435; Best Local Similarity 68.6%; Pred. No. 8.5e-38; Matches 317; Conservative 0; Mismatches 145; Indels 0; Gaps 0; Qy 92 CGCCCTTTCCCTGGGGCGCTGGTGC CGGTGACCGCTGTCCTGCTGCTGCTGCTGCTG 151 Db 355 CACTGTTCCCGCGCGCTGCTGGCGGGCGTCACTGCCACCTTCGGCGCTCTTCGTGG 414 Qy 152 TCGGGGTGAGCGCAACCTGCTGACCTGATGCTGATCGGGCGCTACCGGACATCGGA 211 Db 415 TGGGCATCTCGGGCAACCTGCTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474 Qy 212 CCACCAACCACTTGTACCTGGGCGAGCATGGCCCTGTCGACCTTACTCATCTCTGCTGG 271 Db 475 CCACCAACCACTTACTTATCCATATCAGCATGGCCCTTCTCCGATCTGCTCATCTTCTGTGCA 534 Qy 272 TGGCGTTGACCTGTACCGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGCTCT 331 Db 535 TGGCGTTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594 Qy 332 GCGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391 Db 595 GCAAACTCTTCCAGTGTGTCAGCGAGTGTGTCACCTACCGCAGCTGCTTCCATCACC 654 Qy 392 CGCTAGCGCTGAGCGCTTACTTGGCCATCTGCGCCCGCTGCGCGCGCTGCGCGCGCTTGTGCTCA 451 Db 655 CGCTAGCGCTGAGCGCTTACTTGGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCA 714 Qy 452 CCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG 511 Db 715 CCAAGGGCGTGTGAAGTGGTTCATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774 Qy 512 GTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553 Db 775 GGCCTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 816	
COMMENT	RESULT 2 BF603623 LOCUS DEFINITION ACCESSION VERSION BF603623 269181 MARC 3BOV Bos taurus cDNA 5', mRNA sequence. BF603623 BF603623.1 GI:11701421 EST 25-APR-2001		

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. 608

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A430103P18"

/tissue_type="thymus"

/dev_stage="0 day neonate"

/lab_hosts="DH10B"

/clone_lib="RIKEN full-length enriched, 0 day neonate thymus"

/notes="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATTAATCCCTCCCTCCCTCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 81 a 206 c 183 g 137 t 1 others

ORIGIN

Query Match 17.7%; Score 219.2; DB 14; Length 608;

Best Local Similarity 70.3%; Pred. No. 9.2e-36;

Matches 293; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 137 TGTGCTTGTCTGCTCGGGTGACGGACGACGTGTGACCGGTATGCTGATCGGGCGCT 196

DB 2 TGGCGCTCTTCTGTTGGGATCTCTCGGGCAACCTGTCTCACCATGCTGGTGTGTCGCGCT 61

QY 197 ACCGGGACATCGGACACACCACTTTGTACCTTGGGACAGTGGCGGTGTCGACCTAC 256

DB 62 TCAGGAGCTGGGACACACCACTTACCTTATCCAGATGCGCTTCTCGATCTGC 121

QY 257 TCATCTGCTCGGCTCGGCTTGAACCTGTACCGCTCTTGGCGGTTCGCGGCTTGGGTGT 316

DB 122 TCATCTCTGTGATCGCGCTGACCTCTGCGGCTCTGCGAGTATCGGCTTGAACCT 181

QY 317 TCGGGCGCTGCTGCGGCTGTCTCTACGTGGGCGAGGCTGACCTACGACCGC 376

DB 182 TCGGCGACCTGCTCTGAAACTCTTCCAGTTTGTTCAGCGAGAGTGCACCTACGCCACGG 241

QY 377 TGTGACATGACCGGCTCAGGTCGAGCGCTACCTGGCCATCTGCGCGCTCGCGC 436

DB 242 TCCTACCATACCGGCTGAGGTCGAGGCTACTTCGCCATCTGTTCCCGCTGCGGG 301

QY 437 CCGCGCTTGTGGTACCGCGCGCGGTCCGCGGCTCATCGTGTGCTCTGCGCGCTGG 496

Db 302 CCAAGGTGGTGGTCCACCAAGGCCGTGTGAAGCTGGTCACTCTTGTCAATTTGGGCGGTGG 361

QY 497 CGTGTCTCTCTGCGGTCCTCTTGTCTCTGTTGGGCGTCGAGCAGACCCGCGCA 553

Db 362 CTTTCTGACGCGCGGCGCCATCTTCTGCTGTTGGGCGTGGAGCACGAGACGGCA 418

RESULT 4

BU553576

LOCUS 843 bp mRNA linear EST 16-SEP-2002

DEFINITION AGENCOURT 10242213 NIH_MGC 109 Homo sapiens cDNA clone

IMAGS:6577973 5', mRNA sequence.

ACCESSION BU553576

VERSION BU553576.1 GI:22903848

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 843)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLC2778 row: p column: 05
High quality sequence stop: 534.

FEATURES

source

1. 843

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6577973"

/tissue_type="teratocarcinoma, cell line"

/lab_hosts="DH10B (phage-resistant)"

/clone_lib="NIH MGC 109"

/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC library."

BASE COUNT 213 a 169 c 212 g 247 t 2 others

ORIGIN

Query Match 15.0%; Score 186.4; DB 13; Length 843;

Best Local Similarity 91.0%; Pred. No. 7.4e-29;

Matches 243; Conservative 0; Mismatches 17; Indels 7; Gaps 4;

QY 906 GGTGGTTCTGGCATTTATATAATTTGCTGGTTGCCCTTCCACGTTGGCAGAAATCAATTTACAT 965

DB 534 GGTGGTTCTGGCATTTATATAATTTGCTGGTTGCCCTTCCACGTTGGCAGAAATCAATTTACAT 593

QY 966 AAACACGGAAGATTCGGGATGATGCTCTCTCAGTACTTTAACTCTGCTCTGCA 1025

DB 594 AAACACGGAAGATTCGGGATGATGCTCTCTCAGTACTTTAACTCTGCTCTGCA 653

QY 1026 ACTTTTCTATCTCAGCGCATCTATCAACCAATCTCTACAACTC-ATTTCAAAGAAGT 1084

DB 654 ACTTTTCTATCTCAGCGCATCTATCAACCAATCTCTACAACTC-ATTTCAAAGAAGT 713

QY 1085 AC-AGAGCGCGCGCTTTAAA-----CTGTGCTCGCAAGGAAGTCCAGGCCGAG-AGGCT 1138

DB 714 ACAAGAGCGCGCGCTTTAAAACACTGCTGCTTCGCAAGGGAAGTCCAGGCCGAGAGGCT 773

```

Qy 1139 TCCACAGAGCAGGACACACTCGCGGGG 1165
Db 774 CCCACAGAGCAGGACACTCGCGGGG 800

RESULT 5
BU568940/c
LOCUS BU568940 849 bp mRNA linear EST 16-SEP-2002
DEFINITION AGENCOURT 10400272 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6616239
5', mRNA sequence.
ACCESSION BU568940
VERSION BU568940.1 GI:22919240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 849)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2859 row: j column: 15
High quality sequence stop: 429.
FEATURES
Location/Qualifiers
1..849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6616239"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_82"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgccg); Site_2: SfiI (ggccatcatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCGGCAGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
BASE COUNT 199 a 229 c 193 g 228 t
ORIGIN

Query Match 14.0%; Score 173.4; DB 13; Length 849;
Best Local Similarity 88.6%; Pred. No. 3.9e-26;
Matches 233; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

Qy 981 CGCGATGATGACTTCTCTAGTACTTTAA-CATCGTGGCTCTG--CAACTTTTCTATCT 1037
Db 698 GGGGAAGAGGCTTTTCTCTAGTACTTTAACTTGTGGCACTTTTTTTTGT 639

Qy 1038 GAGCGCATCTATCAACCAATCTCTCAACCTCATTTCAAGAGTACAGAG-CGGCGG 1096
Db 638 AGGGGCATTTATCAACCAATCTTTACCTTCATTTCAAGAGTACAGAGCGCGCGG 579

Qy 1097 CTTTAAACTCTCTCGCAAGGAAGTTCAGGCCGAGAGGCTTCCACAGAGCAGGAC 1156
Db 578 CTTTAAACTCTTCTCGCAAGGAAGTTCAGGCGAGAGGCTTCCACAGAGCAGGAC 520

Qy 1157 CTGCGGGGAAGTTGCGAGGGGACACTGAGGAGACACGGTGGGTACACCGAGACAGCG 1216
Db 519 CTGCGGGGAAGTTGCGAGGGGACACTGAGGGAGACACGGTGGGTACACCGAGACAGCG 460

```

```

Qy 1217 CTAACGTGAAGACGATCGGATAA 1239
Db 459 CTAACGTGAAGACGATCGGATAA 437

RESULT 6
AZ451922/c
LOCUS AZ451922 504 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0251C06R Mouse 10kb plasmid UGUC1M library Mus musculus genomic
clone UGUC1M0251C06 R, genomic survey sequence.
ACCESSION AZ451922
VERSION AZ451922.1 GI:10608187
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 504)
REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
AUTHORS Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0251 row: C column: 06
Seq primer: CACACGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 504.
FEATURES
Location/Qualifiers
1..504
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="UGUC1M0251C06"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGUC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 156 a 110 c 107 g 131 t
ORIGIN

Query Match 11.4%; Score 141.6; DB 28; Length 504;
Best Local Similarity 78.4%; Pred. No. 1.6e-19;

```

Matches 182; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 899 TGCTGGTGGTGGTCTGCATTTATAATTTGC-TGGTTGCCCTTCCACGTTGGCAGATC 957
 Db 252 TGGTGGTGGAGTTCTGGCATTATAGTTTGGTTGGTTTCCAGTTGGCAGATC 193
 QY 958 ATTTACATAAACACGGAGATTCGCGGATGATGTAATCTCTCAGTACTTTTAACATCGTC 1017
 Db 192 TTTTTCATGAACACTGAATAATTCGGGTAATGCACTCTCTCGATATCTTTAATATGGTC 133
 QY 1018 GCTCTGCACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTCAACCTCAATTTCA 1077
 Db 132 ACTCTCAACTTCTATACCGAGTGCATCCATCAAGCCCATCTCCCAACCTCATGTC 73
 QY 1078 AAGAAGTACAGAGCGGGCGCTTTAAACTGCTGCTCCAGAGGAGTCCAGGC 1129
 Db 72 AATTAGTACAAAGCAGTGTGAGGAACTGTTCTGCGGAGACAGTCCAGGC 21

RESULT 7
 LOCUS BI757121 855 bp mRNA linear EST 25-SEP-2001
 DEFINITION 6030882F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200868 5',
 mRNA sequence.
 ACCESSION BI757121
 VERSION BI757121.1 GI:15748699
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strauberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM1503 row: d column: 21
 High quality sequence stop: 803.
 Location/Qualifiers
 1..855
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5200868"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH MGC Library."
 103 a 300 c 299 g 153 t

BASE COUNT
 ORIGIN

Query Match 10.3%; Score 127; DB 12; Length 855;
 Best Local Similarity 55.2%; Pred. No. 2e-16;
 Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGCGCGGAGCGCCGCGTCCGCGCTGCGACGAGCGCG 86
 Db 60 CAGCAGCCCGCGCGCGCGCGCGAGTCTCCAAACCGCGGCTGAGCCGCGCGCT 119

QY 87 CTGCTGCCCCCTTTCCCTGGGGCGCTGGTGCACGCGCTGTGTGTCCTGTGCTGTT 146
 Db 120 GGGGGTGGACACTCGCTCTGGGCCCAAGGTGCTGTTCCACGCGCTCTACGCACTCATCTG 179
 QY 147 CGTCTGCGGGGTGAGCGGCAACGTG-----GTGACCGTGATGCTGATCGGGGCTTACCG 200
 Db 180 GGGCGTGGGCGCGCGGCAATGCGCTGTCGCTGCACGTGGTCTGAAGCGCGGGCGG 239
 QY 201 GGCATGCGGACACCAACCACTTGTACCTGGGCGATGCGGCTGTCGACCTACTCAT 260
 Db 240 GCGCGGGGGCGCTGCGCCACCACTGCTCAGCTTGGCGCTCGCGGCGCTGCTGCTGCT 299
 QY 261 CTGCTCGGGCTCGCGTTGACCTTACCGCTCTGCGCGCTCGCGG---CCCTGGGTGTT 317
 Db 300 GCTGTGCGGGTGGCGTGGAGCTCTACAGTTTGTGTTCCACTACCCCTGGGTCTT 359
 QY 318 CGGGCGCTGCTGCGCGCTGCTTACAGTGGGAGGGCTGACCTACGACGCT 377
 Db 360 CGCGACCTGGGCTGCGCGCTACTTCTGTCACAGCTGTGGCGCTTACGCGCGT 419
 QY 378 GCTGCATGACCGCTCAGCTGAGCGCTACCTGGCCATCTGCGCGCGCTCGCGC 437
 Db 420 GCTGAGCGTGGCGGCTGAGCGCGCTGCTAGCGCTGTGCGCGCTTACGCGCGT 479
 QY 438 CCGCGCTTGTGTCACCGCGCGCTGCGCGCTCATCGCTGTGCTCTGGGCGCTGGC 497
 Db 480 CCGAGCGTGTGACGCGCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 QY 498 GCTGCTCTGCGCGTCCCTTCTGTTCTGTTGCGGCGTGCAGACGACCGCG 550
 Db 540 GCTCGGCTCGCGCTGCGCGCTGCGCGCTCATCGCTGCGCGCGAGACGAACTCG 592

RESULT 8
 LOCUS BM925480 1010 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT 6625013 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5763267
 5', mRNA sequence.
 ACCESSION BM925480
 VERSION BM925480.1 GI:19375859
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strauberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM12814 row: n column: 04
 High quality sequence stop: 703.
 Location/Qualifiers
 1..1010
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5763267"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH MGC Library."
 103 a 300 c 299 g 153 t

BASE COUNT
 ORIGIN

Query Match 10.3%; Score 127; DB 12; Length 855;
 Best Local Similarity 55.2%; Pred. No. 2e-16;
 Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGCGCGGAGCGCCGCGTCCGCGCTGCGACGAGCGCG 86
 Db 60 CAGCAGCCCGCGCGCGCGCGAGTCTCCAAACCGCGGCTGAGCCGCGCGCT 119

upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH MGC Library."

```
BASE COUNT      149 a      348 c      337 g      176 t
ORIGIN
Query Match      10.3%; Score 127; DB 12; Length 1010;
Best Local Similarity 55.2%; Pred. No. 2.1e-16;
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;
QY 27 CGGCCCCGAGGGGGCGGAGCGCGCGTGGCCGCGCTCCCGCTTGCGACGAGCGCG 86
DB 82 CAGCAGCCCGCGCGCCCGCGGCGCCAGCTCAACCCGGGCTGAGCCCTGACGCGCGGCT 141
QY 87 CTGCTCGCCCTTTCCTCGGGGCGCTGGTGGCGGCTGACCGCTGTGCGCTGCTGTT 146
DB 142 GGGCTGGACACTGCGCTCTGGGCGCAAGTGTCTTCCCGCGCTTACGCACTCATCTG 201
QY 147 CGTCTGGGGTGAAGCGCAACGTG-----GTACCGTGATGCTGATCGGGGCGTACCG 200
DB 202 GCGCTGGGCGCGCGGCAATGGCTGTCCGTGCAGTGTGTCTGAAGCGCGGGCGG 261
QY 201 GGACATCGGACACACACCACTTTGACCTGGGCGAGCATGGCGTGTGCGACCTACTCAT 260
DB 262 GCGCGGGGCGCGCTGCGCCACACAGTGTCTGAGCTGGCGCTCGCGGCGCTGCTGCT 321
QY 261 CTGCTCGGGCTCGCGTTCACCTGTACCGCTTGGCGCTCGCG-----CCCTGGGT 317
DB 322 GCTGGTGGCGGTGCGGTGAGCTCTACAGCTTCGTGTGTTCCACTACCCCTGGGTCT 381
QY 318 CGGCGCGCTCTCTGCGCGCTGCTCCCTCTACGTGGGCGAGGCTGACCTACGCCACGCT 377
DB 382 CGCGACCTGGGTGCGCGGCTACTACTTCTGTGCAGAGCTGTGCGCTTACGCCAGGT 441
QY 378 GCTGCAATACCGCTCAGCTGAGCGCTACCTGGCCATCTGCGCGCGCTCGCGCG 437
DB 442 GCTGAGCGTGGCAGGCTGAGCGCGAGCGCTGCTAGCGGTGTGCCAGCCCTGGGTGC 501
QY 438 CCGGCTTTGGTACCGCGCGCGCGCTGCGGGGCTCATCGTGTCTGCGGCGGTGGC 497
DB 502 CCGCAGCTCTCAGCGCACCGCGGACCGCGGTGGTGGCGCTCTCTGTGGGCGCGCTC 561
QY 498 GCTGCTCTCTGCGCGTCCCTTCTTGTCTCTGTTGGGCGTGCAGCAGGACCCCG 550
DB 562 GCTCGGCTCGCCCTGCCATGSCCGTCTATCATGGGCGACAGCAGCACTCG 614
```

```
RESULT 9
BG714306      634 bp      mRNA      linear      EST 08-MAY-2001
LOCUS      60269812F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792730 5',
DEFINITION      mRNA sequence.
ACCESSION      BG714306.1 GI:13993237
VERSION      1 (bases 1 to 634)
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 634)
NTH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-i@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
```

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10671 row: g column: 03
High quality sequence stop: 633.

FEATURES
source

```
1..634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4792730"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2-3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
```

```
BASE COUNT      74 a      221 c      243 g      96 t
ORIGIN
Query Match      10.1%; Score 125.4; DB 10; Length 634;
Best Local Similarity 55.0%; Pred. No. 4.1e-16;
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;
QY 27 CGGCCCCGAGGGGGCGGAGCGCGCGTGGCCGCGCTGCCCGCTTGCGACGAGCGCG 86
DB 49 CAGCAGCCCGCGCGCCCGCGGCGCCAGCTCAACCCGGGCTGAGCTTGAACCGCGGCT 108
QY 87 CTGCTCGCCCTTTCCTCGGGGCGCTGGTGGCGGCTGACCGTGTGCGCTGCTGTT 146
DB 109 GGGCTGGACACTGCGCTCTGGGCGCAAGTGTCTTCCCGCGCTTACGCACTCATCTG 168
QY 147 CGTCTGCGGGTGAAGCGCAACGTG-----GTGACCGTGATGCTGATCGGCGCTACCG 200
DB 169 GCGCTGGGCGCGCGGCGCAATGGCTGTCCGTGCAGCTGGTGTGAAGCGCGGGCGG 228
QY 201 GGACATCGGAGCACACCACTTTGACCTGGGCGAGCATGGCGTGTGCGACCTACTCAT 260
DB 229 GCGAGCGGGCGCGCTGCGGCACACACGTCTGAGCTGGGCGCTCGCGGCGCTGCTGCT 288
QY 261 CTGCTCGGGCTGCGGTTCGACCTGTACCGCTTGGCGCTCGCG-----CCCTGGGT 317
DB 289 GCTGGTGGGCGTGGCGGTGAGCTCTACAGCTTCTGTGTCTTCCACTACCCCTGGGTCT 348
QY 318 CGGCGCGCTCTCTGCGCGCTGTCCCTCTACGTGGGCGAGGCGTGCACCTTACGCCACGCT 377
DB 349 CGCGACCTGGGTGCGCGGCTACTACTTCTGTGCAGAGCTGTGCGCTTACGCCAGGT 408
QY 378 GCTGCAATACCGCGCTCAGCGCTGAGCGCTTACCTGGGCGATCTGCGCGCGCTCGCGCG 437
DB 409 GCTGAGCTGGCAGGCTGAGCGCGCGAGCGCTGCTTAGCGGTGTGCGCAGCCCTGGGTGC 468
QY 438 CCGGCTTTGGTCAACCGCGCGCGCTCGCGGCGCTCATCGTGTGCTCTGCGGCGGTGGC 497
DB 469 CCGCAGCTCTGTCAGCACCCACCGCGGACCGCGGTGGCTGGTGGCGCTCTCTGTGGGCGCGCTC 528
QY 498 GCTGCTCTCTGCGCGTCCCTTCTTGTCTCTGGTGGGCGGTGCGAGCAGGACCCCG 550
DB 529 GCTCGGCTCGCCATGCCCCATGCGCCCTCATCATGGGCGAGACGAGCACTCG 581
```

```
RESULT 10
BM808829
LOCUS      6617918 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734336
DEFINITION      5', mRNA sequence.
ACCESSION      BM808829
VERSION      BM808829.1 GI:19125652
```

```
BM808829      1055 bp      mRNA      linear      EST 05-MAR-2002
AGENCOURT_6617918 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734336
5', mRNA sequence.
BM808829
BM808829.1 GI:19125652
```

[illegible]


```
Qy 176 CCGTGATGCTGATCGGGCGCTACCGGACATGCGGACACCACTTGTACCTGGGCA 235
Db 186 TGTGCTGGTGTCTGCGACGACGAGCTATGAAGACGCGCCACCACTTACCTTTCA 245
Qy 236 GCATGGCGGTGTCGACCTACTATCTGCTC-----GGCTGCGCTTGCACCTGTACCGCC 292
Db 246 GCCTGGCGGTCTCTGACCTCTCTGCTCTCTGCTCTCTTGAATGCGGCTGTATGAGA 305
Qy 293 TCTGGCGCTGCGGCGCTTGGGTGTTGGGCGGCTGCTGCGGCTGTGCTCTACGTGG 352
Db 306 TGTGGCGCAACTACCTTTCTTGTTCGGGCCGCTGGGCTGTACTTCAAGACGCCCTCT 365
Qy 353 GCGAGGGCTGCACCTAGCCACGCTGTGTCATGACCGGCTCAGCGGTGAGGCGTACC 412
Db 366 TTGAGACCGTGTGCTTCGCTCCATCTCTCAGCATCACCACTGAGGCGCTACG 425
Qy 413 TGGCATCTGCGGCGGCTGCGGCGGCTGCTGCTGTCACCGCGGCGGCTGCGGCGG 472
Db 426 TGGCATCTACACCGGCTGCGGCGGCTGTCAGAGCAGCAGCGGCGGCGGCGGCTCAGGA 485
Qy 473 TCATCGCTGTGCTCTGGG 490
Db 486 TCCTGGCATGCTCTGGG 503

RESULT 12
BM543497 1027 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION AGENCOURT_6492823 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726591
5', mRNA sequence.
ACCESSION BM543497
VERSION BM543497.1 GI:18773939
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL12719 row: e column: 24
High quality sequence stop: 635.
Location/Qualifiers
1. 1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5726591"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 124"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcorV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dr primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 153 a 357 c 337 g 180 t
ORIGIN
```

```
Query Match 10.1%; Score 124.8; DB 12; Length 1027;
Best Local Similarity 56.0%; Pred. No. 6.2e-16;
Matches 281; Conservative 0; Mismatches 212; Indels 9; Gaps 2;

Qy 58 CCGCGCTGCGGCTTGGGACGAGCGCGTGTCTCGCCCTTCCCTTGGGGGCGCTGGTG 117
Db 83 CACCCGGGCTGAGCCTGGACGCCCGGCTGGGCTGGGCACTCGCTCTCTGGGCAAGTG 142
Qy 118 CCGGTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171
Db 143 CTGTTACCGGCTCTACGACATCATCTGGGCGCTGGGCGGCGGCAATGGCTGTCC 202
Qy 172 GTACCGGTGATGTGATTCGGGCGCTACCGGGAATGCGGACCAACCACTTGTACCTG 231
Db 203 GTGACGCTGTGTGTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 262
Qy 232 GCAGCATGCGGCGCTGTCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 291
Db 263 AGCTGCGGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
Qy 292 CTCTGGGCGCTGCGGCGG---CCCTGGGCTGTTGGGCGGCTGCTGCTGCTGCTGCTG 348
Db 323 TTGCTGTGTTCCACTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382
Qy 349 GTGGGCGAGGCTGCACCTACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 408
Db 383 GTGACGAGCTGTGCGCTACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
Qy 409 TACTTGGGCTGCTGCGGCGGCTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
Db 443 TGCTTACCGCTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502
Qy 469 GCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
Db 503 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562
Qy 529 GTGGGCGTTCGAGCAGGACCCCG 550
Db 563 ATGGGCGAGAAGACGACGAACTCG 584

RESULT 13
BM546496 1144 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION AGENCOURT_6494007 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5729257
5', mRNA sequence.
ACCESSION BM546496
VERSION BM546496.1 GI:18779503
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL12726 row: e column: 02
High quality sequence start: 27
High quality sequence stop: 519.
Location/Qualifiers
1. 1144
/organism="Homo sapiens"
/mol_type="mRNA"
```



```
Qy 501 GCTCTGCGCGTCCCTTCTTGTCTTGTGGGGTGGAGCAGGACCCCGGCACTCCGT 560
    |||
Db 127 GCTGCGCGTGCACACAGCGTGCACGGCATCGAGGSGTGGCGCCCAAGTCCGC 68
    |||
Qy 561 AGTCCCGGGCTCAATGGCAGCGCGGATCGCTCTCGCTCTCGCTCGTGG 615
    |||
Db 67 CGGAGTTTCTCTCGCGCATCTGCTGTAGGTCCCTCGAACCTCTGGTCTCG 13

RESULT 15
BM808823
LOCUS BM808823 1603 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT 6617870 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734333
5', mRNA sequence.
ACCESSION BM808823
VERSION BM808823.1 GI:19125646
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1603)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM12739 row: h column: 14
High quality sequence start: 84
High quality sequence stop: 514.
Location/Qualifiers
1. 1603
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5734333"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site_1: EcorV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 317 a 579 c 433 g 266 t 8 others
ORIGIN

Query Match 9.6%; Score 119; DB 12; Length 1603;
Best Local Similarity 54.5%; Pred. No. 1.1e-14;
Matches 286; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

Qy 27 CGGCGCGCGGCGCGGAGCGCGCTGCGCGCGCTTGGACGAGCGCG 86
    |||
Db 131 CAGAGCGCGCGCGCGCGCGCGCTCAACCCGGGGCTGAGCTGGACGCCCGCT 190
    |||
Qy 87 CTGCTCGCGCTTCCCTCGGGGCGCTGCTGCGCGCTGCTGTGCTGTGCTGTT 146
    |||
Db 191 GGGCGTGGACACTCGCTCTGGGCAAGGTGCTGTTACCGCGCTCTAGGCACATCTG 250
    |||
Qy 147 CGTCTCGGGGTGAGCGGCAACGTG-----GTACCGTGTATGCTGCGGCGCTACCG 200
    |||
Db 251 GGCCTGGGCGCGGCGGCAATGCGCTGTCCGTGCACGTGTGTGAAGCGCGGCGCG 310
    |||
```

```
Qy 201 GGAATGCGGACCAACCAACTTGTACTTGGGAGCATGGCCGTGTCCGACCTACTCAT 260
    |||
Db 311 GCGCGGCGGCGCGCTGCGGCACACAGTGTCTCAGCTTGGCGCTCGCGGCTGTGCTGCT 370
    |||
Qy 261 CTTGCTCGGGCTGCGGTTGACCTGTACCGCTC---TGGCGGTGCGGGGCTTGGGTGTT 317
    |||
Db 371 GCTGGTGGGCTGCGGTTGGAGCTCTACAGCTTCTGTGTGTTCCACTTACCCCTGGGTCTT 430
    |||
Qy 318 CGGCGCGCTGCTCTGCGCGCTGCTTACGTGGGAGGGGTGACACTTACGCCACGCT 377
    |||
Db 431 CGGGGACCTGGGCTGCGCGGCTACTTCTGTGTCAGAGCTGTGCGGCTTACGCCACGCT 490
    |||
Qy 378 GCTGCACATGACCGCGCTCAGCGTCGAGCGCTACCTGGCCATCTGCGCGCGCTTCCGCGC 437
    |||
Db 491 GCTGAGCGTGGCAGGCTGAGCGCGCGAGCGCTGCTAGCGGTGTGCCAGCCCTGCGCTGC 550
    |||
Qy 438 CCGGCTTGGTCAACCGGCGCGCGCTCGCGGCGCTCATCGCTGTGCTGTGGGCGGTGGC 497
    |||
Db 551 CCGGAGCTGCTGACGCGCACCGCGGACCCCGTGGTGGCGCTCTCTGTGGGCGGCTC 610
    |||
Qy 498 GCTGCTCTGCGCGTCCCTTCTTGTCTTGTGGGCGCTCGAGCA 542
    |||
Db 611 GCTCGGCTCCACCTGCGCCATGGCCGACATCATGGGCGCAGAAGCA 655
    |||
```

Search completed: January 1, 2004, 02:52:40
Job time : 2913.48 secs

THIS PAGE BLANK (usptb)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 52.1454 Seconds
(without alignments)
1254.097 Million cell updates/sec

Title: US-09-719-485-3

Perfect score: 2155

Sequence: 1 MGSPWNSGDPGEGAREPPWP.....DTGGDTVGYTETSANYKTMG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_19Jun03.*

```

1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	2155	100.0	412 21	AA54145 Amino acid sequenc
2	2155	100.0	412 22	AA62652 Long form of motil
3	2155	100.0	412 22	AA68476 Amino acid sequenc
4	2155	100.0	412 23	ABG30936 Human G protein-co
5	2155	100.0	412 24	ABP81993 Human motilin rece
6	2155	100.0	501 22	AA65822 Human GPR38 varian
7	2149	99.7	412 21	AA62854 Human G protein co
8	2141	99.4	412 23	ABB09535 Human motilin rece
9	1747	81.1	400 22	AA68477 Amino acid sequenc

10	1581	73.4	386	21	AA54146	Amino acid sequenc
11	1581	73.4	386	22	AA62653	Short form of motil
12	1149.5	53.3	271	22	AA68476	Amino acid sequenc
13	907	42.1	363	21	AA54147	The puffer fish mo
14	907	42.1	363	22	AA68479	Amino acid sequenc
15	865.5	40.2	349	21	AA69293	A canine growth ho
16	863.5	40.1	366	21	AA90832	Human G protein-co
17	863.5	40.1	366	21	AA90345	Human G protein-co
18	863.5	40.1	366	22	AA97376	Rat growth hormone
19	863.5	40.1	366	22	AA62650	Human G-protein co
20	863.5	40.1	366	23	ABB09534	Human ghrelin rece
21	863	40.0	353	18	AAW1908	Pig growth hormone
22	863	40.0	353	18	AAW19215	Swine growth hormo
23	861.5	40.0	364	21	AA54565	A mouse growth hor
24	861.5	40.0	364	22	AA97377	Rat growth hormone
25	860	39.9	361	18	AAW19217	Human growth hormo
26	860	39.9	362	18	AAW19610	Human growth hormo
27	857.5	39.8	364	18	AAW19613	Rat growth hormone
28	857.5	39.8	366	21	AA90866	Human mutant G pro
29	845.5	39.2	364	18	AAW19220	Rat growth hormone
30	674.5	31.3	271	18	AAW19612	Human growth hormo
31	674.5	31.3	271	18	AAW19219	Human growth hormo
32	644.5	29.9	289	18	AAW19611	Human growth hormo
33	644.5	29.9	289	24	ABP81828	Human growth hormo
34	641	29.7	289	18	AAW19609	Pig growth hormone
35	639.5	29.7	289	18	AAW19216	Swine growth hormo
36	639.5	29.7	289	18	AAW19218	Human growth hormo
37	501	23.2	418	17	AA98562	Human neurotensin
38	497	23.1	418	24	ABP81861	Human neurotensin
39	497	23.1	542	22	ABG12316	Novel human diagno
40	491	22.8	418	22	ABP56371	Non-endogenous hum
41	486.5	22.6	445	22	AAU68523	Human novel cytol
42	486.5	22.6	445	22	AAU68566	Human novel cytol
43	485.5	22.5	403	21	AA90638	Human G protein-co
44	485.5	22.5	403	21	AA44642	Human growth hormo
45	485.5	22.5	403	22	AA899185	Human FM-3. Homo

ALIGNMENTS

RESULT 1

AA54145

ID AA54145 standard; Protein; 412 AA.

XX AC AA54145;

XX AC AA54145;

XX DT 27-MAR-2000 (first entry)

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI Pong S, Smith RG;
XX
XX WPI; 2000-105868/09.
DR N-PSDB; AAZ45403.
XX
PT Novel receptor protein for screening compounds used in treating
PT irritable bowel syndrome, constipation and other gastric conditions -
XX
PS Claim 3; Fig 3; 44pp; English.
XX
CC The present sequence represents splice variant MTL-R1A of the motilin
CC receptor. The gene encodes a G-protein coupled receptor, and is
CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
CC functional seven transmembrane domain form, and MTL-R1B is a truncated
CC five transmembrane domain. The MTL-R1 proteins are used to identify
CC agonists and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathic pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile colic, irritable
CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX
SQ Sequence 412 AA;
Query Match 100.0%; Score 2155; DB 21; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.2e-194;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPALPPCDERRCSPFPLGALVPTAVTAVCLFVVGSGNVTV 60
DB 1 MGSPWNGSDGPEGAREPPALPPCDERRCSPFPLGALVPTAVTAVCLFVVGSGNVTV 60
QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYRLMRSRPFGLICRLSLYVGE 120
DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYRLMRSRPFGLICRLSLYVGE 120
QY 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLAVALLSAGPFLVGE 180
DB 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLAVALLSAGPFLVGE 180
QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240
DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240
QY 241 QLGNALRVMLVTTAYFFLPFLCLISILYGLIRELWSSRPRLRGPASGRGRHQTVRVL 300
DB 241 QLGNALRVMLVTTAYFFLPFLCLISILYGLIRELWSSRPRLRGPASGRGRHQTVRVL 300
QY 301 LVVLAFLICWLPFHVGRIIYNTEDSRMVFYSYFNIVALQLFYLASINPLINLSK 360
DB 301 LVVLAFLICWLPFHVGRIIYNTEDSRMVFYSYFNIVALQLFYLASINPLINLSK 360
QY 361 KYRAAFLKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKTWG 412
DB 361 KYRAAFLKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKTWG 412
RESULT 2
AAB62652
ID AAB62652 standard; Protein; 412 AA.
XX
AC AAB62652;
XX
XX 23-JUL-2001 (first entry)
XX
DE Long form of motilin receptor, GPR-38A isoform.

XX zsig33; signal transduction; hormone; enzyme; neural development;
KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
KW glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R;
KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.
XX
OS Homo sapiens.
XX
XX WO200138355-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US32074.
XX
XX 22-NOV-1999; 99US-0166765.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
XX
XX WPI; 2001-355879/37.
XX N-PSDB; AAF83683.
XX
XX Forming reversible peptide receptor complex for purifying cell and
XX peptides, stimulating signal transduction and modulating hormone
XX secretion, involves contacting a receptor with zsig33 polypeptide -
XX
XX Disclosure; Page 102-104; 11pp; English.
XX
XX The invention relates to a method of forming a reversible peptide-
XX receptor complex that involves providing an immobilized receptor, and
XX contacting the receptor with a zsig33 peptide (comprising residues 24-37
XX of AAB62649), where the receptor binds to the zsig33 peptide. The method
XX is useful for purifying cells, purifying a peptide, stimulating signal
XX transduction in a cell expressing a receptor. It is also useful for
XX modulating secretion of hormones, neural development and/or utilization,
XX gastric contractility, nutrient uptake, secretion of digestive and
XX pancreatic enzymes and hormones, secretion of insulin-like growth factor
XX -I, secretion of non-zsig33 proteins. It is useful for modulating growth
XX hormone secretion in a mammal having a disease associated with abnormal
XX levels of growth hormone, such as osteoporosis, bone repair, bone
XX remodeling, low osteoblast levels, cartilage repair and remodeling,
XX skeletal dysplasia, immune suppression, obesity, growth retardation,
XX protein catabolic responses after surgery, cachexia, protein loss,
XX dwarfism, wound healing and ovulation induction, treating a mammal having
XX a metabolic disorder requiring neurological feedback, such as satiety
XX regulation, glucose absorption and metabolism and neuropathy-associated
XX gastrointestinal disorders, and stimulating glucose-induced insulin
XX release in a mammal. The present sequence represents the long form of
XX motilin receptor, GPR-38A (one of the two isoforms of GPR38 which result
XX from alternative splicing). GPR38 has homology to the human G-protein
XX coupled receptor, GHS-R.
XX
SQ Sequence 412 AA;
Query Match 100.0%; Score 2155; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.2e-194;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPALPPCDERRCSPFPLGALVPTAVTAVCLFVVGSGNVTV 60
DB 1 MGSPWNGSDGPEGAREPPALPPCDERRCSPFPLGALVPTAVTAVCLFVVGSGNVTV 60
QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYRLMRSRPFGLICRLSLYVGE 120
DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYRLMRSRPFGLICRLSLYVGE 120
QY 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLAVALLSAGPFLVGE 180
DB 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLAVALLSAGPFLVGE 180
QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240

```

|||||
Db 181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPGPETAEEAALFSRECRPSA 240
QY 241 QLGALRVMLVMTTAYFPFLPCLLSILYGLIGRELMSSRRPLRGPAAASGRGRHRTQTVRVL 300
Db 241 QLGALRVMLVMTTAYFPFLPCLLSILYGLIGRELMSSRRPLRGPAAASGRGRHRTQTVRVL 300
QY 301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
Db 301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
QY 361 KYRAAFKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412
Db 361 KYRAAFKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412

RESULT 3
AAB68478
ID AAB68478 standard; Protein; 412 AA.
XX
AC AAB68478;
XX
DT 23-JUL-2001 (first entry)
XX
DE Amino acid sequence of a human motilin receptor polypeptide.
XX
KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea.
XX
OS Homo sapiens.
XX
PN WO200132710-A1.
XX
PD 10-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US29426.
XX
PR 29-OCT-1999; 99US-0162264.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Tan C, McKee K;
XX
DR WPI; 2001-343479/36.
DR N-PSDB; AAF85449.
XX
PT Novel polypeptides related to dog and rabbit motilin receptor
PT polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX
PS Disclosure; Page 32-33; 42pp; English.
XX
CC The present sequence represents a human motilin receptor polypeptide.
CC The specification describes a unique sequence present in exon 1 of
CC the dog motilin receptor, which is not present in human or Sphearoides
CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.
XX
SQ Sequence 412 AA;
Query Match 100.0%; Score 2155; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.2e-194;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPCDERRCSPFPLGALVPVTAVCLCLFVVGVSNGVVTY 60
Db 1 MGSPWNGSDGPEGAREPPWALPCDERRCSPFPLGALVPVTAVCLCLFVVGVSNGVVTY 60

```

```

QY 61 MLIGRYRDMRTTNTNLYLGSMAVSDLLILGLPPDLVRLWRSRPMWFGPPLCRLSLYYVGE 120
Db 61 MLIGRYRDMRTTNTNLYLGSMAVSDLLILGLPPDLVRLWRSRPMWFGPPLCRLSLYYVGE 120
QY 121 CTYATLLHTALTALSVRYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPPLPLVGVE 180
Db 121 CTYATLLHTALTALSVRYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPPLPLVGVE 180
QY 181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPGPETAEEAALFSRECRPSA 240
Db 181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPGPETAEEAALFSRECRPSA 240
QY 241 QLGALRVMLVMTTAYFPFLPCLLSILYGLIGRELMSSRRPLRGPAAASGRGRHRTQTVRVL 300
Db 241 QLGALRVMLVMTTAYFPFLPCLLSILYGLIGRELMSSRRPLRGPAAASGRGRHRTQTVRVL 300
QY 301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
Db 301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
QY 361 KYRAAFKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412
Db 361 KYRAAFKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412

RESULT 4
ABG30936
ID ABG30936 standard; Protein; 412 AA.
XX
AC ABG30936;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human G protein-coupled receptor 38 (GPR38).
XX
KW Human; G protein-coupled receptor 38; receptor; GPR38;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;
KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX
OS Homo sapiens.
XX
FH 15..31 Location/Qualifiers
FT Region /note= "Antigenic fragment"
FT Region 217..232 /note= "Antigenic fragment"
FT Region 276..291 /note= "Antigenic fragment"
FT Region 373..388 /note= "Antigenic fragment"
FT Region /note= "Antigenic fragment"
XX
WO200257791-A2.
XX
PN 25-JUL-2002.
XX
PF 29-NOV-2001; 2001WO-US45219.
XX
PR 29-NOV-2000; 2000US-250251P.
PR 30-NOV-2000; 2000US-250452P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Brown JP, Burner GC, Roush CL, Kulander BG;
XX WPI: 2002-566812/60.
DR N-PSDB; ABK90112.
XX
PT Assay for detecting Alzheimer's disease, Parkinson's disease,
PT ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
PT carcinoma, comprises using a binding partner for G protein coupled

```

PT receptor 38 -
 PS Disclosure; Fig 2; 112pp; English.
 XX
 CC The present invention relates to a new assay method that involves
 CC contacting a binding partner specific for G protein coupled receptor
 CC (GPR) 38 with specific cells. The method of the invention is useful for
 CC the detection of an increased risk of Alzheimer's disease, Parkinson's
 CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
 CC glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
 CC for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
 CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
 CC glioblastoma, breast carcinoma, colon carcinoma, lung small cell
 CC carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
 CC pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
 CC to manufacture a medicament able to reduce the symptoms of these
 CC diseases. Nucleic acids encoding GPR 38 can also be used to treat the
 CC diseases. The present amino acid sequence represents the human G
 CC protein-coupled receptor 38 (GPR38) of the invention.
 XX
 SQ Sequence 412 AA;
 Query Match 100.0%; Score 2155; DB 23; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.2e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPPLGALVPTAVCLCLFVVGVSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPPLGALVPTAVCLCLFVVGVSGNVTV 60
 QY 61 MLIGRYDMRTTTNLYLGSMVSDLLILGLPFDLYRLWSRPWVFGPFLCRSLYVGE 120
 DB 61 MLIGRYDMRTTTNLYLGSMVSDLLILGLPFDLYRLWSRPWVFGPFLCRSLYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGE 180
 DB 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGE 180
 QY 181 QDPGISVVGGLNGTARTASSPLASSPPLWLSRAPPPSPGPTAAEALFSRCPSPA 240
 DB 181 QDPGISVVGGLNGTARTASSPLASSPPLWLSRAPPPSPGPTAAEALFSRCPSPA 240
 QY 241 QLGLALVLMVWTAYFFLPFLCLISLYLGRLWSRRPLRGPAAASGRGRHQTVRVL 300
 DB 241 QLGLALVLMVWTAYFFLPFLCLISLYLGRLWSRRPLRGPAAASGRGRHQTVRVL 300
 QY 301 LVVVLAFIICWLPFHVGRIIYINTDSRMVYFQYFNIVALQLFYLSASINPILYNLISK 360
 DB 301 LVVVLAFIICWLPFHVGRIIYINTDSRMVYFQYFNIVALQLFYLSASINPILYNLISK 360
 QY 361 KYRAAAFKLLARKSRPRGHRSDTAGVAGDTGDTVGYTETSANVKTGM 412
 DB 361 KYRAAAFKLLARKSRPRGHRSDTAGVAGDTGDTVGYTETSANVKTGM 412

RESULT 5

ABP81993
 ID ABP81993 standard; Protein; 412 AA.
 XX
 AC ABP81993;
 XX
 DT 04-MAR-2003 (first entry)
 XX
 DE Human motilin receptor GPR38 protein SEQ ID NO:473.
 XX
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 Query Match 100.0%; Score 2155; DB 24; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.2e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPPLGALVPTAVCLCLFVVGVSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPPLGALVPTAVCLCLFVVGVSGNVTV 60
 QY 61 MLIGRYDMRTTTNLYLGSMVSDLLILGLPFDLYRLWSRPWVFGPFLCRSLYVGE 120
 DB 61 MLIGRYDMRTTTNLYLGSMVSDLLILGLPFDLYRLWSRPWVFGPFLCRSLYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGE 180
 DB 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGE 180
 QY 181 QDPGISVVGGLNGTARTASSPLASSPPLWLSRAPPPSPGPTAAEALFSRCPSPA 240

KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer.
 XX Homo sapiens.
 OS
 PN WO200261087-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 19-DEC-2001; 2001WO-US50107.
 XX
 PR 19-DEC-2000; 2000US-257144P.
 XX
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX
 PI Burmer GC, Roush CL, Brown JP;
 XX WPI; 2003-046718/04.
 DR N-PSDB; ABZ42842.
 XX
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating
 PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
 PT cancer or autoimmune diseases -
 XX Disclosure; Fig 1; 523pp; English.
 PS
 XX The present invention describes antigenic peptides (I) comprising:
 CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular
 CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity
 CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and
 CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
 CC an antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention.

Db 181 QDPGISVVGGLNGTARIASSPLASPLMLSRAPPSPGPTAEAAALFSRECRPSA 240
 Qy 241 QLGALRVMLWTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAAASGRERGHRTQTVRVL 300
 Db 241 QLGALRVMLWTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAAASGRERGHRTQTVRVL 300
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 360
 Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 360
 Qy 361 KYRAAFLKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 412
 Db 361 KYRAAFLKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 412

RESULT 6
 AAG65822
 ID AAG65822 standard; Protein; 501 AA.
 XX AC AAG65822;
 DT 30-JAN-2002 (first entry)
 XX DE Human GPR38 variant GPR38V polypeptide.
 XX KW GPR38V; variant; antibacterial; cytostatic; analgesic; antiasthmatic;
 KW anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;
 KW antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;
 KW antiulcer; antiemetic; cardiac; vaccine; human.
 XX OS Homo sapiens.
 XX PN WO200164836-A2.
 XX PD 07-SEP-2001.
 XX PF 28-FEB-2001; 2001WO-US06277.
 XX PR 01-MAR-2000; 2000US-0516315.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Elshourbagy N, Shabon U;
 XX DR WPI; 2001-638956/73.
 XX DR N-PSDB; AAI65989.
 XX PT New human GPR38V polypeptide and polynucleotide, useful for treating
 PT e.g. bacterial, fungal, protozoal and viral infections, cancers or
 PT allergies, as vaccines, and for identifying agonists and antagonists
 PT potentially useful in therapy -
 XX FS Claim 1; Page 26; 32pp; English.
 XX CC This represents a human GPR38 variant (GPR38V) polypeptide. GPR38V can
 CC be expressed by standard recombinant methodology. The polynucleotides and
 CC polypeptides are used in the treatment of bacterial, fungal, protozoal
 CC and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,
 CC diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart
 CC failure, hypertension, urinary retentions, osteoporosis, allergies,
 CC ulcers, migraine, psychotic and neurological disorders, or dyskinesias.
 CC They are also useful for identifying agonists and antagonists that are
 CC potentially useful in therapy, as vaccines to induce immunological
 CC response in a mammal. The polypeptides may also be used as immunogens to
 CC produce antibodies immunospecific for the polypeptides, and to identify
 CC membrane bound or soluble receptors.
 XX SQ Sequence 501 AA;

Query Match 100.0%; Score 2155; DB 22; Length 501;
 Best Local Similarity 100.0%; Pred. No. 1.5e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSPFWNSDGPGEARBPWPALPPCDERRCSPPFLGALVPVTAVCCLCFVVGSGNVVTV 60
 Db 90 MGSPFWNSDGPGEARBPWPALPPCDERRCSPPFLGALVPVTAVCCLCFVVGSGNVVTV 149
 Qy 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWPGPLICRLSLYVGE 120
 Db 150 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWPGPLICRLSLYVGE 209
 Qy 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLIAVLWAVALLSAGPPLFLVGV 180
 Db 210 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLIAVLWAVALLSAGPPLFLVGV 269
 Qy 181 QDPGISVVGGLNGTARIASSPLASPLMLSRAPPSPGPTAEAAALFSRECRPSA 240
 Db 270 QDPGISVVGGLNGTARIASSPLASPLMLSRAPPSPGPTAEAAALFSRECRPSA 329
 Qy 241 QLGALRVMLWTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAAASGRERGHRTQTVRVL 300
 Db 330 QLGALRVMLWTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAAASGRERGHRTQTVRVL 389
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 360
 Db 390 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 449
 Qy 361 KYRAAFLKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 412
 Db 450 KYRAAFLKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 501

Human G protein coupled receptor hGPR38 (V297K) protein SEQ ID NO:130.
 Human; G protein coupled receptor; GPCR; transmembrane receptor;
 identification; agonist; screening; therapeutic; pharmaceutical;
 mutant.
 Homo sapiens.
 Synthetic.
 WO200022131-A2.
 20-APR-2000.
 13-OCT-1999; 99WO-US24065.
 13-OCT-1998; 98US-0170496.
 12-NOV-1998; 98US-0108029.
 20-NOV-1998; 98US-0109213.
 27-NOV-1998; 98US-0110060.
 16-FEB-1999; 99US-0120416.
 26-FEB-1999; 99US-0121852.
 12-MAR-1999; 99US-0123944.
 12-MAR-1999; 99US-0123945.
 12-MAR-1999; 99US-0123946.
 12-MAR-1999; 99US-0123948.
 12-MAR-1999; 99US-0123949.
 12-MAR-1999; 99US-0123951.
 28-MAY-1999; 99US-0136436.
 28-MAY-1999; 99US-0136437.
 28-MAY-1999; 99US-0136439.
 28-MAY-1999; 99US-0137127.
 28-MAY-1999; 99US-0137131.
 30-JUN-1999; 99US-0141448.
 27-AUG-1999; 99US-0151114.

PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX
 DR WPI; 2000-317986/27.
 DR N-PSDE; AAA46116.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX
 PS Example 2; Page 168-169; 187pp; English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (Orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX

SQ Sequence 412 AA;

Query Match 99.7%; Score 2149; DB 21; Length 412;
 Best Local Similarity 99.8%; Pred. No. 4.5e-194;
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Qy 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLLPDLYRLWRSRPPWFGPLCLRLSLVYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLLPDLYRLWRSRPPWFGPLCLRLSLVYVGE 120
 Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRAIALVAVALLSAGPFLVGVGE 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRAIALVAVALLSAGPFLVGVGE 180
 Qy 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240
 Qy 241 QLGALRVMLWVTYAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Db 241 QLGALRVMLWVTYAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360
 Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360
 Qy 361 KYRAAFAKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTYSANVKTWG 412
 Db 361 KYRAAFAKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTYSANVKTWG 412

RESULT 8

ID ABB09535
 ID ABB09535 standard; protein; 412 AA.

XX AC ABB09535;

XX DT 22-OCT-2002 (first entry)

XX DE Human motilin receptor.

XX

KW Human; motilin receptor; appetite; food intake; agonist; analogue;
 KW undernutrition; anorexia; cachexia; malignant disease; infection;
 KW inflammatory disease; weight loss; antagonist; obesity; anorectic;
 KW anabolic; ghrelin receptor homologue; receptor.
 XX
 OS Homo sapiens.
 XX
 PN WO200260472-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-JP00765.
 XX
 PR 31-JAN-2001; 2001JP-0024423.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Inui A, Asakawa A, Kaga T;
 XX
 DR WPI; 2002-619206/66.
 XX
 PT Remedies for diseases with hypo-nutrition status e.g. inappetence and
 PT cachexia, containing ghrelin or its analogs including agonists and
 PT antagonists -
 XX
 PS Disclosure; Fig 1B; 50pp; Japanese.
 XX
 CC The invention relates to the use of ghrelin or its analogues for the
 CC treatment of diseases associated with undernutrition such as anorexia,
 CC and also relates to the use of ghrelin antagonists for the prevention
 CC or treatment of obesity. The invention additionally discloses a method
 CC for screening ghrelin agonists or antagonists by measuring the amount
 CC of food intake, neuropeptide Y (NPY) expression, binding of NPY to
 CC NPY receptor Y1, oxygen consumption, gastric emptying, or activity of
 CC the vagus nerve. Intracerebroventricular (ICV) administration of ghrelin
 CC in animals was found to increase food intake over a period of 24 hours.
 CC Ghrelin and its analogues may therefore be used to treat conditions
 CC such as loss of appetite, anorexia, cachexia, malignant diseases, and
 CC weight loss associated with infection or inflammatory diseases.
 CC Conversely, ghrelin antagonists may be used in the treatment of obesity.
 CC The present sequence represents the human motilin receptor, a homologue
 CC of the ghrelin receptor (ABB09534) which is referred to in the disclosure
 CC of the invention.
 XX
 SQ Sequence 412 AA;

Query Match 99.4%; Score 2141; DB 23; Length 412;
 Best Local Similarity 99.5%; Pred. No. 2.5e-193;
 Matches 410; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Qy 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLLPDLYRLWRSRPPWFGPLCLRLSLVYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLLPDLYRLWRSRPPWFGPLCLRLSLVYVGE 120
 Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRAIALVAVALLSAGPFLVGVGE 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRAIALVAVALLSAGPFLVGVGE 180
 Qy 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240
 Qy 241 QLGALRVMLWVTYAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Db 241 QLGALRVMLWVTYAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360
 Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360

Qy 361 KYRAAFKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKTWG 412
 Db 361 KYRAAFKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKTWG 412

RESULT 9
 AAB68477
 ID AAB68477 standard; Protein; 400 AA.
 AC AAB68477;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Amino acid sequence of a rabbit motilin receptor polypeptide.
 XX
 KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO200132710-A1.
 XX
 PD 10-MAY-2001.
 XX
 PF 25-OCT-2000; 2000WO-US29426.
 XX
 PR 29-OCT-1999; 99US-0162264.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Tan C, McKee K;
 XX
 DR WPI; 2001-343479/36.
 DR N-PSDB; AAF85448.
 XX
 PT Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans -
 XX
 PS Claim 8; Page 18; 42pp; English.
 XX

The present sequence represents a rabbit motilin receptor polypeptide.
 The specification describes a unique sequence present in exon 1 of
 the dog motilin receptor, which is not present in human or Sphaeroides
 nepheus 75E7 motilin receptor sequences. The unique nucleic acid
 sequence is useful for measuring the ability of a compound to affect
 motilin receptor activity. Motilin receptor polynucleotides and
 polypeptides are used to identify therapeutic compounds which are
 useful for treating gastrointestinal diseases and disorders such as
 gastric motility disorders, gastroparesis, irritable bowel syndrome,
 and diarrhoea.

Query Match 81.1%; Score 1747; DB 22; Length 400;
 Best Local Similarity 83.9%; Pred. No. 3.5e-156;
 Matches 344; Conservative 11; Mismatches 43; Indels 12; Gaps 3;

Qy 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPFFLGALVPVTAIVCLCLFVVGSGNVVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPFFLGALVPVTAIVCLCLFVVGSGNVVTV 60

Qy 61 MLIGRYDMRTTNYLGSMVSDLLILGLPFDLYRLWRSPRVFGLLCRLSLYVGE 120
 Db 61 LLIGRYDMRTTNYLGSMVSDLLILGLPFDLYRLWRSPRVFGLLCRLSLYVGE 120

Qy 121 CTYATLLHMTALSVRYLAICRPARVLTTRRRVRLIAVLVALLSAGPEFLVGV 180
 Db 121 CTYASLHMTALSVRYLAICRPARVLTTRRRVRLIAVLVALLSAGPEFLVGV 180

Qy 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPPPSGPTAEAAALFSPRECRPSA 240

Db 181 QDPGAVFAADPNRGTV-----PLDPSP-----APASPPSGP-GAFAALFSPRECRPSA 228
 Qy 241 QLGALRVMLWVTTAYFPFLCLSLIYGLIGRELWSRRPLRGPAASGRGRHRTVRL 300
 Db 229 QLGALLRVMLWVTTAYFPFLCLSLIYGLIARGLWRGRLGPAATGRGRHRTVRL 288
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTDSRMVPSQYFNIVALQFLYLSASINPILYNLISK 360
 Db 289 LVVLAFLIICWLPFHVGRIIYINTDSRMVPSQYFNIVALQFLYLSASINPILYNLISK 348
 Qy 361 KYRAAFKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKT 410
 Db 349 KYRAAARLLRBSRAGSGVCGSRGPEQDVAGDTGGDTAGCTTSTANTKT 398

RESULT 10
 AAY54146
 ID AAY54146 standard; Protein; 386 AA.
 XX
 AC AAY54146;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Amino acid sequence of the motilin receptor splice variant MTL-R1B.
 XX
 KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation.
 XX
 OS Homo sapiens.
 XX
 PN WO9964436-A1.
 XX
 PD 16-DEC-1999.
 XX
 PF 08-JUN-1999; 99WO-US12773.
 XX
 PR 12-JUN-1998; 98US-0089098.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;
 XX
 DR WPI; 2000-105868/09.
 DR N-PSDB; AAZ45404.
 XX

Novel receptor protein for screening compounds used in treating
 irritable bowel syndrome, constipation and other gastric conditions -
 Claim 5; Fig 5; 44pp; English.

The present sequence represents splice variant MTL-R1B of the motilin
 receptor. The gene encodes a G-protein coupled receptor, and is
 designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 functional seven transmembrane domain form, and MTL-R1B is a truncated
 five transmembrane domain. The MTL-R1 proteins are used to identify
 agonists and antagonists which can be used for treating gastric motility
 disorders, functional defects, disorders secondary to neurological
 disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 dysmotility, diabetes, infections, stress-related motility disorders,
 psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 constipation, chronic idiopathic pseudo obstruction, acute faecal
 impaction, postoperative ileus, gallstones, infantile colic, irritable

CC bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.

XX
 SQ Sequence 386 AA;

Query Match 73.4%; Score 1581; DB 21; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.5e-140;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGSGNVTV 60
 Db 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGSGNVTV 60

Qy 61 MLIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWFGLCLRLSLYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWFGLCLRLSLYVGE 120

Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLMAVALLSAGPFLVGV 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLMAVALLSAGPFLVGV 180

Qy 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240

Qy 241 QLGLARVLMVLTWTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTVRVL 300
 Db 241 QLGLARVLMVLTWTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTVRVL 300

RESULT 11
 AAB62653
 ID AAB62653 standard; Protein; 386 AA.

XX
 AC AAB62653;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Short form of motilin receptor, GPR-38B isoform.

XX zsig33; signal transduction; hormone; enzyme; neural development;
 KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 KW glucose; osteopathic; anorectic; vulnertary; immunomodulator; GHS-R;
 KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.

XX
 OS Homo sapiens.
 XX
 XX WO200138355-A2.
 XX
 XX 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32074.
 XX
 XX 22-NOV-1999; 99US-0166765.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
 XX
 XX WPI; 2001-355879/37.
 XX
 XX N-PSDB; AAF83684.
 XX
 XX Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide -
 XX
 XX Disclosure; Page 106-109; 11pp; English.

XX
 CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37

CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the short form of
 CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result
 CC from alternative splicing). GPR38 has homology to the human G-protein
 CC coupled receptor, GHS-R.

XX
 SQ Sequence 386 AA;

Query Match 73.4%; Score 1581; DB 22; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.5e-140;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGSGNVTV 60
 Db 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGSGNVTV 60

Qy 61 MLIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWFGLCLRLSLYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWFGLCLRLSLYVGE 120

Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLMAVALLSAGPFLVGV 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLMAVALLSAGPFLVGV 180

Qy 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240

Qy 241 QLGLARVLMVLTWTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTVRVL 300
 Db 241 QLGLARVLMVLTWTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTVRVL 300

RESULT 12
 AAB68476
 ID AAB68476 standard; Protein; 271 AA.

XX
 AC AAB68476;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Amino acid sequence of a dog motilin receptor exon 1.

XX
 KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 XX
 XX Canis sp.
 XX
 XX WO200132710-A1.
 XX
 XX 10-MAY-2001.
 XX
 XX 25-OCT-2000; 2000WO-US29426.
 XX
 XX 29-OCT-1999; 99US-0162264.
 XX
 XX (MERI) MERCK & CO INC.

Db	196	MWVSSVFFFLPVFCLTVLYGLIGRKLWRRGRGDTAGGASLREQSHRQTVKMLAVVVPF	255
Qy	308	IIICWLPFHVGRRIIYINTEDS---RMVYFSQYFNIVALQLFYLSASINPILYNLISKYRA	364
Db	256	IFCWLPHVGRYLFPSKSFPGSLEIAQISQYCNLVSVFLFYLSAANPILYNIMSKYRV	315
Qy	365	AAFKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGYTETSAN	407
Db	316	AVFKLLGFEPFSQRKSLTKDESSR-----AWTESSIN	348

Search completed: January 1, 2004, 06:46:43
Job time : 54.1454 Secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:22:35 ; Search time 17.5539 Seconds
(without alignments)
993.061 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDGPEGAREPPWP.....DTGGDTVGYTETSANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	863.5	40.1	366	3	US-09-077-675A-13
2	863.5	40.1	366	4	US-09-077-674-13
3	863.5	40.1	366	4	US-09-170-496D-88
4	863	40.0	353	3	US-09-077-675A-3
5	863	40.0	353	4	US-09-077-674-3
6	861.5	40.0	364	3	US-09-077-675A-16
7	861.5	40.0	364	4	US-09-077-674-16
8	860	39.9	361	3	US-09-077-675A-8
9	860	39.9	361	4	US-09-077-674-8
10	857.5	39.8	366	4	US-09-170-496D-210
11	775.5	36.0	302	3	US-09-077-675A-2
12	775.5	36.0	302	4	US-09-077-674-2
13	774.5	35.9	302	3	US-09-077-675A-7
14	774.5	35.9	302	4	US-09-077-674-7
15	674.5	31.3	271	3	US-09-077-675A-12
16	674.5	31.3	271	4	US-09-077-674-12
17	644.5	29.9	289	3	US-09-077-675A-10
18	644.5	29.9	289	4	US-09-077-674-10
19	633	29.4	289	3	US-09-077-675A-5
20	633	29.4	289	4	US-09-077-674-5
21	485.5	22.5	403	4	US-09-170-496D-114
22	483.5	22.4	403	4	US-09-170-496D-224
23	479	22.2	415	4	US-09-545-944-2
24	466	21.6	353	1	US-08-118-270-45
25	466	21.6	353	5	PCT-US93-08528-45
26	413	19.2	410	3	US-08-858-876A-2
27	413	19.2	410	3	US-09-472-880-2

Query Match 40.1%; Score 863.5; DB 3; Length 366;
Best Local Similarity 44.7%; Pred. No. 1.6e-66;

ALIGNMENTS

RESULT 1
US-09-077-675A-13
; Sequence 13, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-13

Sequence 4, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 8, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli

Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAVCICLCLFVVGSGN 56
DB 2 WNATPSEEPFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVALFVVGIA 61
QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSL 116
DB 62 LLTMLVVSFRFRELRTTNNLYLSSMAFSDLLIFLCMPDLVLRLWQYRPMWFGDLLCKLFQ 121
QY 117 VGEQCTVATLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPPLFL 176
DB 122 VSESCYATVLTITLALSVRYFAICFPLRAKVVVTKGRVKLVFVIMAVAFCSAGPTFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPSPPSGPTAEAAALFSRECR 236
DB 182 VGVHE-----NGT-----DP-W-----DTNECR 199
QY 237 PS--PAQALGALRVMLVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRGHR 294
DB 200 PTEFAVSGLLTVWVSSIFFPLPVCLVLYSLIGRLWRRRGDAVVVGASLRDQNHK 259
QY 295 QTVRVLLVVLAFIICWLPFHVGRIIYINTEDS---RMMYFSQYFNIVALQLFVLSASIN 351
DB 260 QTVKMLAVVVFATILCWLPHVGRYLFSPGSLBIAQISQYCNLVSFVLYLSAAN 319
QY 352 PILYNIISKYRAAFKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
DB 320 PILYNIISKYRVAVFRLGFPFSPQKSLTKDESSR-----AWTESSIN 365

RESULT 2

US-09-077-674-13
; Sequence 13, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077, 674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-13

Query Match 40.1%; Score 863.5; DB 4; Length 366;
Best Local Similarity 44.7%; Pred. No. 1.6e-66;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAVCICLCLFVVGSGN 56
DB 2 WNATPSEEPFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVALFVVGIA 61
QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSL 116
DB 62 LLTMLVVSFRFRELRTTNNLYLSSMAFSDLLIFLCMPDLVLRLWQYRPMWFGDLLCKLFQ 121
QY 117 VGEQCTVATLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPPLFL 176
DB 122 VSESCYATVLTITLALSVRYFAICFPLRAKVVVTKGRVKLVFVIMAVAFCSAGPTFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPSPPSGPTAEAAALFSRECR 236
DB 182 VGVHE-----NGT-----DP-W-----DTNECR 199
QY 237 PS--PAQALGALRVMLVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRGHR 294
DB 200 PTEFAVSGLLTVWVSSIFFPLPVCLVLYSLIGRLWRRRGDAVVVGASLRDQNHK 259
QY 295 QTVRVLLVVLAFIICWLPFHVGRIIYINTEDS---RMMYFSQYFNIVALQLFVLSASIN 351
DB 260 QTVKMLAVVVFATILCWLPHVGRYLFSPGSLBIAQISQYCNLVSFVLYLSAAN 319
QY 352 PILYNIISKYRAAFKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
DB 320 PILYNIISKYRVAVFRLGFPFSPQKSLTKDESSR-----AWTESSIN 365

RESULT 3

US-09-170-496D-88
; Sequence 88, Application US/09170496D
; Patent No. 6553339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-170-496D-88

Query Match 40.1%; Score 863.5; DB 4; Length 366;
Best Local Similarity 44.7%; Pred. No. 1.6e-66;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAVCICLCLFVVGSGN 56
DB 2 WNATPSEEPFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVALFVVGIA 61
QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSL 116
DB 62 LLTMLVVSFRFRELRTTNNLYLSSMAFSDLLIFLCMPDLVLRLWQYRPMWFGDLLCKLFQ 121

Qy	117	VGECTTATLHMTALSVERYLACRPIRARVVLVTRRRVRALIAVLWAVALLSAGPFIFL	176
Db	122	VSECTYATLTITALSVERYFAICFPRAKVVVTGKGLVFIVFWAFVACSAGPFIWL	181
Qy	177	VGVDQPGISVVPGLNGTGARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECR	236
Db	182	VGVEHE-----NCT-+-----DP-W-----DTNECR	199
Qy	237	PS--PAQIGALRVMLWTTAYFFFLCCLSLTYLGILGRELWSSRRRLRGPAASGRGHHR	294
Db	200	PTEFARVSGLLTVMVWSSIIFFFLPVFCFLTLYLSLIGRKLWRRRRGDVAVGSASLRDQNHK	259
Qy	295	QTFRVLLVVLAEIICMLPHFGRTIYYINTEDS---RMMYSOYNENIVALQLFYLSASN	351
Db	260	QTVKMLAVVFPAILCHLPFHVGRYLFPSKSPGPSLEIAQISOYCNLVSVFLFYLSAAIN	319
Qy	352	PILNYLSKKYRAAAFLLLARKSRPRGFHSRDDTAGEVAGDTGGDTVGYTETSAN	407
Db	320	PILYNIIMSKYRVAVFRLLGPEFPSPQRKLSTLKDESSR-----AWTESIN	365

RESULT 4
US-09-077-675A-3
; Sequence 3, Application US/09077675A

```

; Patent NO. 62421199
;
; GENERAL INFORMATION:
;
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
;
; TITLE OF INVENTION: RECEPTOR ASSAY
;
; NUMBER OF SEQUENCES: 16
;
; CORRESPONDENCE ADDRESS:
; ADDRESS: Merck & Co., Inc.
; STREET: P. O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
;
; COUNTRY: USA
;
; ZIP: 07065-0900
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
;
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: DOS
;
; SOFTWARE: FASTSEQ for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/097,675A
; FILING DATE: 3-JUN-1998

```

; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cocuzzo, Anna L.
 ; REGISTRATION NUMBER: 42,452
 ; REFERENCE/DOCKET NUMBER: 19590P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1273
 ; TELEFAX: 732-594-4720
 ;

```

; SEQUENCE:
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-3

```

Query Match 40.0%; Score 863; DB 3; Length 353;
Best Local Similarity 46.5%; Pred. No. 1.7e-66;
Matches 186; Conservative 53; Mismatches 95; Indels

```

Qy 19 WPALPPCD---ERRCSPPLGALVPVTAVCCLCFVWGNGVNTVMLIGRYDRM
Db 8 WDAPPENDSLVBELLPLFPTLLAGVTATCVALPVGTAIGNLLTMLVVSFRPNR
Qy 76 YLGSNAVSDLLITLLGLPEDLYLWRSRPWFGLCLBLSLYVGGCTYATLLHMT
Db 68 YLSSNAFSDLLITFLCMLJDLFRLMQYRPNWNLGNLLCKLQFVBSCTYATVLTIT
Qy 136 RYLAICRPLURVLVTRRRVRALIAVMAVALLSAGPFLVLGVGEQDPGISVVP
Db 128 RYFAICFPLRAKVVTYKRVKLVILVIAVAFCSAGPIFVLGVGHEH-----
Qy 196 RIASGLASSPPLMUSRAPPPSPGPGTBAALFBSRECRPS---PAQLGALRVMM
Db 178 -----DPRD-----TNECRATEFAVRSGLLTVMM
Qy 254 AVFFELPCLSLYGLIGRELWSSRRPLRGPAASG---RERGRQTVRVLVLLVVL
Db 206 VFFELPVCLTVLYSLIGRKW---RRKRGBAVGSLLRDQNHQTVKMLAVVVF
Qy 311 WLPHVHGRIIY---INTEDSRMVFYSQYFNIVALQLFVLSASINPILYNLISKKY
Db 263 WLPHVHGRIYFSKSLPGSVBIAQISQYCNLVSPVLPVLSAAINPILYNIMSKKY
Qy 368 KULLARKSRPGFHRSRDTAGEVAGDTGGDTVGTETSAN 407
Db 323 KULGPEFPFSQRKSLTKDESSR-----AWTESSIN 352

RESULT 5
US-09-077-674-3
; Sequence 3, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

RESULT 5

US-09-077-674-3
; Sequence 3, Application US/09077674

Patent No. 6531314

: APPLICANT: Arena, Joseph P.
 : APPLICANT: Cully, Doris F.
 : APPLICANT: Feighner, Scott D.
 : APPLICANT: Howard, Andrew D.
 : APPLICANT: Liberator, Paul A.
 : APPLICANT: Schaeffer, James M.
 : APPLICANT: Van der Ploeg, Leon
 : TITLE OF INVENTION: GROWTH HORMONE
 : NUMBER OF SEQUENCES: 16
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Merck & Co., Inc.
 : STREET: P.O. Box 2000, 126 E.

;
; STREET: BOX 2000, 120 E. JEFFERSON AVE
; CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

; COMPUTER READABLE

;
MEDIUM TYPE: Di

; COMPUTER: IBM C

; OPERATING SYSTEM

; SOFTWARE: FastS

;
CURRENT APPLICATION

APPLICANT	APPLICATION NUMBER	STATUS	DATE
...

FILING DATE: 3-3-2010

CLASSIFICATION: PRIOR APPLICATION

APPLICANT'S NAME: _____
 ADDRESS: _____
 CITY: _____ STATE: _____ ZIP: _____
 PHONE: _____
 FAX: _____
 E-MAIL: _____
 DATE: _____

	APPLICANT NAME	APPLICATION NUMBER	FILING DATE
;			
:			

FILING DATE: 11/11/2011
ATTORNEY/AGENT NAME: THE FIRM

ATTORNEY/AGENCY NAME: COCUIZZO

NAME: COCUZZO,
REGISTRATION NUM

REGISTRATION NUMBER
REFERENCE/DOCKET

REFERENCE/ DOCKET
TELECOMMUNICATION

TELEPHONE: 732-

TELEFAX: 732-59

TELEX:

; INFORMATION FOR SEQ

SEQUENCE CHARACTER

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-3

Query Match 40.0%; Score 863; DB 4; Length 353;
Best Local Similarity 46.5%; Pred. No. 1.7e-66;
Matches 186; Conservative 53; Mismatches 95; Indels 66; Gaps 9;

QY 19 WPALPPCD---ERRCSPFPLGALVPVTAVALCLFVGVSGNVVTVMLIGRYDMRTTTL 75
DB 8 WDAPPENDSLVEBLLPLFPPTLLAGVTATCVAFVVGAGNLLTMLVSVRFEMRTTTL 67

QY 76 YLGSMVSDLLIILGLPFDLYRLWRSPWPVFGPDLRLSLYVGECCYATLLHMTALSVE 135
DB 68 YLSSMAFSDLLIFLCMPDLFLRLQVPRPMLNGLLCKLFQVSESCYATVLTITALSVE 127

QY 136 RYLAICRPLARVLRVRRRRLVIAVLAVALISAGPFLVGVGEODPGISVVVPLNGTA 195
DB 128 RYFAICFPLRAKVVTGRVKLVILVIAVAFCSAGPIFVLVGEHD-----NGT- 177

QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS--PAQLGALRVMLVTT 253
DB 178 -----DPRD-----TNECRATEFAVRSGLLTVMVVSS 205

QY 254 AYFPLFCLISILYGLIGRLWSRPLRPAASG---RRGRQTVRVLLVVLAFIIC 310
DB 206 VFFFLFVCLTVLSYLIIGRLW---RRKGEAAVGSRLRDQNHQVTKMLAVVVFILC 262

QY 311 WLPFHVGRIIY---INTEDSRMYFQYFNIVALQLFYLSASINPILYNIISKYRAAAF 367
DB 263 KLLGFPFQSKRLSTLKDESSR---SQYCNLVSFVLYSAAINPILYNIISKYRVAVF 322

QY 368 KLLARKSRPRGRPHRSTAGEVAGDTGDTGVYTTSETAN 407
DB 323 KLLGFPFQSKRLSTLKDESSR-----AWTESSIN 352

```

```

RESULT 6
US-09-077-675A-16
; Sequence 16, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pal, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.

```

```

; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-16

Query Match 40.0%; Score 861.5; DB 3; Length 364;
Best Local Similarity 48.3%; Pred. No. 2.4e-66;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGPEGAREP-----PWPALPPCD---ERRCSPFPLGALVPVTAVALCLFVGVSG 55
DB 2 WNAT--PSBEPEPNVTLDLDWDASPGNDSLDPDELLPLFPAPLLAGVTATCVAFVVGISG 59

QY 56 NVVTVMILIGRYDMRTTTLNLYGSMVSDLLIILGLPFDLYRLWRSPWPVFGPDLRLSL 115
DB 60 NLLTMLVSVRFELRITTLNLYSSMAFSDLLIFLCMPDLVRLQVPRPMLNGLLCKLFQ 119

QY 116 YVGECCYATLLHMTALSVERYLAICRPLARVLRVRRRRLVIAVLAVALISAGPFLF 175
DB 120 FVSESCYATVLTITALSVERYFAICFPLRAKVVTGRVKLVILVIAVAFCSAGPIFV 179

QY 176 LVCVEODPGISVVVPLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSREC 235
DB 180 LVGEVHE-----NGT-----DPRD-----TNEC 197

QY 236 RPS--PAQLGALRVMLVTTTAYFPLFCLISILYGLIGRLWSRPLRPAASG---RE 290
DB 198 RATEFAVRSGLLTVMVVSSVFFFLFVCLTVLSYLIIGRLW---RR--RQDAVAGSLRD 253

QY 291 RGRQTVRVLLVVLAFIICWLPFHVGRIIYINTEDS---RMVYFQYFNIVALQLFYLS 347
DB 254 QNHQVTKMLAVVVFVAFILCWLPHFVGRVLFSEKSEFSGLEIAQISQYCNLVSFVLYLS 313

QY 348 ASINPILYNIISKYRAAFKLL 370
DB 314 AAINPILYNIISKYRVAVFKLL 336

```

```

RESULT 7
US-09-077-674-16
; Sequence 16, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Joseph P.
; APPLICANT: Arena, Doris F.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

```

; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-16

Query Match 40.0%; Score 861.5; DB 4; Length 364;
Best Local Similarity 48.3%; Pred. No. 2.4e-66;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

Qy 5 WNSDGPGEAREP-----PWPALPPCD---ERRCSPPPLGALVPVTAVCLCLFVVGVS 55
Db 2 WNAT--PSEBPEPNVTLDDWDASPGNDSLPLDELLPLFPAPLLAGVTATCVALFVVGIS 59

Qy 56 NVVTMLIGRYDRMTTNNLYLGSMAVSDLLILGLPDLVRLWRSPWVFGPLLCRLSL 115
Db 60 NLLTMLVVSFRRLTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDLCKLQF 119

Qy 116 YVGECTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLAVLWAVALLSAGPFLF 175
Db 120 FVSESTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVILVWAVAFCSAGPIFV 179

Qy 176 LVGVEQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPSGTETAALFRRRC 235
Db 180 LVGVEH-----NGT-----DPRD-----TNEC 197

Qy 236 RPS--PAQLGALRVMLVTTAYFPFLPCLSLYLIGRELWSSRRPLRGPAAGS---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFLPVCLTVLYSLIGRKLW--RR--RGDAVAGSLRD 253

Qy 291 RGHRTVRLVLLVVLAFIICWLPVHVGRIIYINTEDS---RMVFSQYFNIVALQLFVLS 347
Db 254 QNHKQTVKMLAVVVFVAFILCWLPFHVGRYLFKSPFPGSLEIAQISQYCNLVSVFLFVLS 313

Qy 348 ASINPILYNIISKYRAAFKLL 370
Db 314 AAINPILYNIIMSKYRVAVFKLL 336

RESULT 8
US-09-077-675A-8
; Sequence 8, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.

; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-8

Query Match 39.9%; Score 860; DB 3; Length 361;
Best Local Similarity 46.1%; Pred. No. 3.2e-66;
Matches 183; Conservative 57; Mismatches 97; Indels 60; Gaps 8;

Qy 19 WPALPPCD---ERRCSPPPLGALVPVTAVCLCLFVVGVSNNVTVMLIGRYDRMTTNNL 75
Db 16 WDASPGNDSLGLDELQLFPAPLLAGVTATCVALFVVGIAGNLLTMLVSRPRELRTTNNL 75

Qy 76 YLGSMAVSDLLILGLPDLVRLWRSPWVFGPLLCRLSLYVGECTYATLLHMTALSVE 135
Db 76 YLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDLCKLQFVSESTYATVLTITALSVE 135

Qy 136 RYLAICRPLRARVLTTRRRVRLAVLWAVALLSAGPFLFVGVQDPGISVVPGLNGTA 195
Db 136 RYFAICPPLRAKVVTGKRVKLVIFVWAVAFCSAGPIFVLVGVHE-----NGT- 185

Qy 196 RIASSPLASSPPLWLSRAPPPSPSGTETAALFRRRCPS--PAQLGALRVMLVTT 253
Db 186 -----DP--W-----DTNECRTEPAVRSGLLTVMWVSS 213

Qy 254 AYFPLPCLSLYLIGRELWSSRRPLRGPAAGSRRGRHRTVRLVLLVVLAFIICWLP 313
Db 214 IFPFLPVCLTVLYSLIGRKLWRRRGDAVVGASLRDQNHKQTVKMLAVVVFVAFILCWLP 273

Qy 314 FHVGRIIYINTEDS---RMVFSQYFNIVALQLFVLSASINPILYNIISKYRAAFKLL 370
Db 274 FHVGRYLFKSPFPGSLEIAQISQYCNLVSVFLFVLSAANPILYNIIMSKYRVAVFKLL 333

Qy 371 LARKSRPRGPHRRDRTAGEVAGDTGGDTGVGTETSAN 407
Db 334 GFEPFSQKSLTLKDESSR-----AWTESSIN 360

RESULT 9
US-09-077-674-8
; Sequence 8, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
```


STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-2

Query Match 36.0%; Score 775.5; DB 3; Length 302;
Best Local Similarity 46.6%; Pred. No. 4.7e-59;
Matches 166; Conservative 48; Mismatches 79; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLYVGE 119
DB 1 MLVSVRFREMTTNLYLSSMAFSDLLIFLCMPDLFLWQYRPMNLGNLLCKLQFVSE 60
QY 120 GCTYATLLHMTALSVERYLAIICPLRPARVLTTRRRVRLAVLWAVALLSAGPFLVLGV 179
DB 61 SCTYATVLTITALSVERYLAIICPLRPARVLTTRRRVRLAVLWAVALLSAGPFLVLGV 120
QY 180 EQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALPSRECRPS - 238
DB 121 EHD-----NGT-----DPRD-----TNECRATE 138
QY 239 -PAQALGALRVLMTVTYAFPLPCLISLYGLIGRELWSSRRPLRGPAASG---RERGHR 294
DB 139 FAVRSGLLTVMWVSVFPLPCLTVLYSLIGRKLW---RRKGEAAVGSLLRDQNHK 195
QY 295 QTVRVLLVVLVLAIIICWLPFHVGRYIY---INTEDSRMYFSQYFNIVALQLPYLSASIN 351
DB 196 QTVKMLAVVVFALICWLPFHVGRYLSKLEPGSVIEIAQISQYCNLVSFVLYLSAAIN 255
QY 352 PILYNLISKYRAAFAKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
DB 256 PILYNIMSKKYRVAVFKLIGFEPFSQKLSLTKDESSR-----AMTESSIN 301

RESULT 12
US-09-077-674-2
Sequence 2, Application US/09077674
Patent No. 6531314
GENERAL INFORMATION:
APPLICANT: Arena, Joseph P.
APPLICANT: Cully, Doris F.
APPLICANT: Feighner, Scott D.
APPLICANT: Howard, Andrew D.
APPLICANT: Liberator, Paul A.
APPLICANT: Schaeffer, James M.
APPLICANT: Van Der Ploeg, Leonardus

TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: 3-JUN-1998
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19589P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-674-2

Query Match 36.0%; Score 775.5; DB 4; Length 302;
Best Local Similarity 46.6%; Pred. No. 4.7e-59;
Matches 166; Conservative 48; Mismatches 79; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLYVGE 119
DB 1 MLVSVRFREMTTNLYLSSMAFSDLLIFLCMPDLFLWQYRPMNLGNLLCKLQFVSE 60
QY 120 GCTYATLLHMTALSVERYLAIICPLRPARVLTTRRRVRLAVLWAVALLSAGPFLVLGV 179
DB 61 SCTYATVLTITALSVERYLAIICPLRPARVLTTRRRVRLAVLWAVALLSAGPFLVLGV 120
QY 180 EQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALPSRECRPS - 238
DB 121 EHD-----NGT-----DPRD-----TNECRATE 138
QY 239 -PAQALGALRVLMTVTYAFPLPCLISLYGLIGRELWSSRRPLRGPAASG---RERGHR 294
DB 139 FAVRSGLLTVMWVSVFPLPCLTVLYSLIGRKLW---RRKGEAAVGSLLRDQNHK 195
QY 295 QTVRVLLVVLVLAIIICWLPFHVGRYIY---INTEDSRMYFSQYFNIVALQLPYLSASIN 351
DB 196 QTVKMLAVVVFALICWLPFHVGRYLSKLEPGSVIEIAQISQYCNLVSFVLYLSAAIN 255
QY 352 PILYNLISKYRAAFAKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
DB 256 PILYNIMSKKYRVAVFKLIGFEPFSQKLSLTKDESSR-----AMTESSIN 301

RESULT 13
US-09-077-675A-7
Sequence 7, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pat, Lee-Yuh

Db 259 YNIMSKYRVAVFLLGFBPFSQKSLTLKDESSR-----AWTESSIN 301

RESULT 15

US-09-077-675A-12
; Sequence 12, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/077, 675A
; APPLICATION NUMBER: US/09/077, 675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-12

Query Match 31.3%; Score 674.5; DB 3; Length 271;
Best Local Similarity 45.0%; Pred. No. 2e-50;
Matches 145; Conservative 44; Mismatches 76; Indels 57; Gaps 7;
QY 91 LPFDVRLWRSPWVFGPLLCRLSLVYGGCTVATLLHMTALSVERVLAICRPLRARVLV 150
Db 1 MPDLVLRLWQYRPNWFGDLCKLFQVSECTATVLTITLALSVERVFAICFPLRAKVVV 60
QY 151 TRRRVRLALVLAVALLSAGPFLVLGVQDPGISVVPGLNGTARIASSPLASSPPLWL 210
Db 61 TKGRVKLVIFVIMAVAFCSAGPIFVLVGEHE-----NGT-----DP--W- 98
QY 211 SRAPPPSPGPGPETAFAALFSPRECRPS--PAQLGALRVMLWVTYATFFLPFLCLSLY 268
Db 99 -----DTNECRPTFAVRSGLLTVMVWSSIEFFFLPVFCLTVLYS 138
QY 269 LIGRELWSSRRPLRGPAAASGRCHROTVRVLLVWLAFIICWLPFHVGRILIIINTEDS- 327
Db 139 LIGRKLWRRRRGDVAVGASLRDQNHQKTVKMLAVVFAFILCWLFPFHVGRYLFKSFEPG 198
QY 328 --RMVYFSQYFNIVALQLFVLSASINPILYNLISKYRAAAAFKLLLARKSRPRGFHRSRD 385

Db 199 SLEIAQISQYCNLVSFVLFYLSAAINPILYNIMSKKYRVAVFRLGFBPFSQKSLTLKD 258
QY 386 TAGEVAGDTGGDTVGYTETSAN 407
Db 259 ESSR-----AWTESSIN 270

Search completed: January 1, 2004, 06:48:27
Job time : 19.5539 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query #			DB	ID	Description
		Match	Length				
1	2155	100.0	412	12	US-10-318-661-28	Sequence 28, App	
2	2155	100.0	412	12	US-10-206-677-2	Sequence 2, Appl	
3	2155	100.0	412	15	US-10-225-567A-473	Sequence 473, Appl	
4	2155	100.0	412	15	US-10-290-078-15	Sequence 15, App	
5	2149	99.7	412	12	US-10-417-830A-130	Sequence 130, Appl	
6	2149	99.7	412	12	US-10-417-830A-150	Sequence 150, Appl	
7	863.5	40.1	366	12	US-10-303-204A-13	Sequence 13, App	
8	863.5	40.1	366	15	US-10-251-385-88	Sequence 88, App	
9	863	40.0	323	12	US-10-303-204A-3	Sequence 3, Appl	
10	861.5	40.0	364	12	US-10-303-204A-16	Sequence 16, App	
11	860	39.9	361	12	US-10-303-204A-8	Sequence 8, Appl	
12	857.5	39.8	366	15	US-10-251-385-210	Sequence 210, Appl	
13	775.5	36.0	302	12	US-10-303-204A-2	Sequence 2, Appl	
14	774.5	35.9	302	12	US-10-303-204A-7	Sequence 7, Appl	
15	674.5	31.3	271	12	US-10-303-204A-12	Sequence 12, Appl	

[illegible]

	Query Match	100.0%	Score 2155;	DB 12;	Length 412;
	Best Local Similarity	100.0%;	Pred. No. 5.4e-167;		
	Matches 412;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGSPWNGSDGPEGAREPPHPALPPCDERRCSPFPLGALVPTAVCLCLFVVGSGNNVTY	60		
Db	1	MGSPWNGSDGPEGAREPPHPALPPCDERRCSPFPLGALVPTAVCLCLFVVGSGNNVTY	60		
Qy	61	MLIGRYDRMRTTNNLYLGSMVSDLLILLCPLDPDLYRLWRSRPWVGPLLCRLSLYVGEG	120		
Db	61	MLIGRYDRMRTTNNLYLGSMVSDLLILLCPLDPDLYRLWRSRPWVGPLLCRLSLYVGEG	120		

Qy	121	CTYATLLHNTALSVERYLAICPLBARVLVTRRRVRALIAVLMAVALLSAGPFLFVGVE	180
Db	121	CTYATLLHNTALSVERYLAICPLBARVLVTRRRVRALIAVLMAVALLSAGPFLFVGVE	180
Qy	181	QDPGISVPELNGTARIASSPLASSPPPLWSRAPPPSPGPETAEEAAALFSRECRPSPA	240
Db	181	QDPGISVPELNGTARIASSPLASSPPPLWSRAPPPSPGPETAEEAAALFSRECRPSPA	240
Qy	241	QLGALRVMLVWTTAYFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRGHQTVRVL	300
Db	241	QLGALRVMLVWTTAYFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRGHQTVRVL	300
Qy	301	LVVVLAFIICWLPFHVGRITIIYNTEDSRMWFQSYENIVALQLPYLSASINPILYNLISK	360
Db	301	LVVVLAFIICWLPFHVGRITIIYNTEDSRMWFQSYENIVALQLPYLSASINPILYNLISK	360
Qy	361	KYRAAFAKLLARKSPRGFHRSDRTAGEVAGDTGGDTGVGTYSANVKTMG	412
Db	361	KYRAAFAKLLARKSPRGFHRSDRTAGEVAGDTGGDTGVGTYSANVKTMG	412

RESULT 2

US-10-206-677-2
; Sequence 2, Application US/10206677
; Publication NO. US20030186336A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences, Inc.
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; APPLICANT: Kulander, Bruce G.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS RELATED TO
; TITLE OF INVENTION: GPCR 38, A G PROTEIN-COUPLED RECEPTOR (GPCR)
; FILE REFERENCE: 1920-1-8
; CURRENT APPLICATION NUMBER: US/10/206,677
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/250,251
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,452
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-206-677-2

Query Match	100.0%	Score 2155;	DB 12;	Length 412;
Best Local Similarity	100.0%;	Pred. No. 5.4e-167;		
Matches 412;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGSPWNGSDGPEGAREPPWPALPCDERRRCSFPFLGALVPVTA	VCLCLFVVGVSGNVTV	60
Dd	1	MGSPWNGSDGPEGAREPPWPALPCDERRRCSFPFLGALVPVTA	VCLCLFVVGVSGNVTV	60
Qy	61	MLTGRYDMRTTNLYLGSMVAVSLLILGLPFDLYLWRSRPWFGPL	LRLSLYYGEG	120
Dd	61	MLTGRYDMRTTNLYLGSMVAVSLLILGLPFDLYLWRSRPWFGPL	LRLSLYYGEG	120
Qy	121	CTYTALLHMTALSVERYLAICRPURARLVTRRRVRALIAVLMA	VALLSAGPFLLVGVE	180
Dd	121	CTYTALLHMTALSVERYLAICRPURARLVTRRRVRALIAVLMA	VALLSAGPFLLVGVE	180
Qy	181	QDPGISVVPGLNGTARTASSPLASPPWLWSRAPPPPGPETAEAA	AALFSRECRSPSA	240
Dd	181	QDPGISVVPGLNGTARTASSPLASPPWLWSRAPPPPGPETAEAA	AALFSRECRSPSA	240
Qy	241	QLGALRVMLVWTATAYFFLPFLCISLYCLTGRELWSSRRPLRG	PAASGREGHOTVRVL	300
Dd	241	QLGALRVMLVWTATAYFFLPFLCISLYCLTGRELWSSRRPLRG	PAASGREGHOTVRVL	300
Qy	301	LVTVLAFIIICWLPFHVGRIIVYINTEDSRMVMFYFOYNIVALQ	LFYLSASINPIYLNIISK	360

```

Db      301  LVVLAFLIICWLPFHVGRITIIYINTEDSRMMYFSQYFNIVALQLFYLSASINFLYNLISK 360
QY      361  KYRAAAAFKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGVTYTSANVKTMG 412
Db      361  KYRAAAAFKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGVTYTSANVKTMG 412

RESULT 3
US-10-225-567A-473
; Sequence 473, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 473
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-473

```

Query Match	100.0%	Score	2155	DB 15	Length	412
Best Local Similarity	100.0%	Pred. No.	5.4e-167			
Matches	412	Conservative	0	Mismatches	0	Gaps
Qy	1	MGSPWNGSDGPEGAREPPWALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVVTV	60			
Db	1	MGSPWNGSDGPEGAREPPWALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVVTV	60			
Qy	61	MLIGRYDMRTTNLYLIGSMVAVSDLLILLGLPFDLYRLWRSRPVWFGPLLRLSLVYVGG	120			
Db	61	MLIGRYDMRTTNLYLIGSMVAVSDLLILLGLPFDLYRLWRSRPVWFGPLLRLSLVYVGG	120			
Qy	121	CTVATLHMTALSVERYLACRPLRAVLTTRRVRLIIVLWAVALLSAGPFLFLVGVE	180			
Db	121	CTVATLHMTALSVERYLACRPLRAVLTTRRVRLIIVLWAVALLSAGPFLFLVGVE	180			
Qy	181	QDPCISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPSPGPTAEAAAALFSRECRPSPA	240			
Db	181	QDPCISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPSPGPTAEAAAALFSRECRPSPA	240			
Qy	241	QLGALRVMLWVTAYFFLPFLCJLIGRELWSSRRPLRGPAASGRGRGHROTVRVL	300			
Db	241	QLGALRVMLWVTAYFFLPFLCJLIGRELWSSRRPLRGPAASGRGRGHROTVRVL	300			
Qy	301	LVVVLAFLIICWLPHPVGRITVINTEDSRMMVFSQYENIVALQLFVLSASINPILYNLISK	360			
Db	301	LVVVLAFLIICWLPHPVGRITVINTEDSRMMVFSQYENIVALQLFVLSASINPILYNLISK	360			
Qy	361	KYRAAAAPKLLARKSRPRGFHSRSDTAGEVAGDTGGDTVGYTETSANVKTWG	412			
Db	361	KYRAAAAPKLLARKSRPRGFHSRSDTAGEVAGDTGGDTVGYTETSANVKTWG	412			

RESULT 4

US-10-230-078-15
; Sequence 15, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17692 or 58874

```
; FILE REFERENCE: MPI2001-288P1(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-290-078-15

Query Match      100.0%; Score 2155; DB 15; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGSGNVTV 60

QY 61 MLIGRYDRMTTNTNLVLSMAVSDLLILGLPDLVRLWRSRPWVFGPFLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNTNLVLSMAVSDLLILGLPDLVRLWRSRPWVFGPFLCRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180

QY 181 QDPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240
Db 181 QDPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240

QY 241 QLGALRVMLWVTYAYFPFLCCLISILYGLIGRELWSSRRPLRGPAAASGRERHRTVRL 300
Db 241 QLGALRVMLWVTYAYFPFLCCLISILYGLIGRELWSSRRPLRGPAAASGRERHRTVRL 300

QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQOLFYLASINPILYNLISK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQOLFYLASINPILYNLISK 360

QY 361 KYRAAFAKLLARKSRPRGFRHSDRTAGEVAGDTGGDTVGYTETSANVTMG 412
Db 361 KYRAAFAKLLARKSRPRGFRHSDRTAGEVAGDTGGDTVGYTETSANVTMG 412

RESULT 5
US-10-417-820A-130
; Sequence 130, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
```

```
; FILE REFERENCE: MPI2001-288P1(M)
; CURRENT APPLICATION NUMBER: 60/123,945
; CURRENT FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; PRIOR FILING DATE: 1999-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-130

Query Match      99.7%; Score 2149; DB 12; Length 412;
Best Local Similarity 99.8%; Pred. No. 1.7e-166;
Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGSGNVTV 60

QY 61 MLIGRYDRMTTNTNLVLSMAVSDLLILGLPDLVRLWRSRPWVFGPFLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNTNLVLSMAVSDLLILGLPDLVRLWRSRPWVFGPFLCRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180

QY 181 QDPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240
Db 181 QDPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRPSA 240

QY 241 QLGALRVMLWVTYAYFPFLCCLISILYGLIGRELWSSRRPLRGPAAASGRERHRTVRL 300
Db 241 QLGALRVMLWVTYAYFPFLCCLISILYGLIGRELWSSRRPLRGPAAASGRERHRTVRL 300

QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQOLFYLASINPILYNLISK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQOLFYLASINPILYNLISK 360

QY 361 KYRAAFAKLLARKSRPRGFRHSDRTAGEVAGDTGGDTVGYTETSANVTMG 412
Db 361 KYRAAFAKLLARKSRPRGFRHSDRTAGEVAGDTGGDTVGYTETSANVTMG 412

RESULT 6
US-10-417-820A-150
; Sequence 150, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
```

```

; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; PRIOR FILING DATE: 1999-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 150
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-150

Query Match
Best Local Similarity 99.7%; Score 2149; DB 12; Length 412;
Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWPPALPCDERRCSPPLGALVPTAVCLFVGVSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWPPALPCDERRCSPPLGALVPTAVCLFVGVSGNVTV 60
QY 61 MLIGRYDRMTTNNLYLGSMAVSDLLILLGLPDLVRLWRSRPWVFGPDLCLRLSLY 120
Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILLGLPDLVRLWRSRPWVFGPDLCLRLSLY 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARLVTRRRVRLAVLWAVALLSAGPFLVGV 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARLVTRRRVRLAVLWAVALLSAGPFLVGV 180
QY 181 QDPCISVPGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
Db 181 QDPCISVPGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRVMLVTTAYFPLPCILSILYGLIGRELWSSRRPLRGPASGRGHRQTVRL 300
Db 241 QLGALRVMLVTTAYFPLPCILSILYGLIGRELWSSRRPLRGPASGRGHRQTVRL 300
QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQFLYLSASINPILYNLSK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQFLYLSASINPILYNLSK 360
QY 361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTGYTETSANVKTMG 412
Db 361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTGYTETSANVKTMG 412

```

RESULT 7

```

US-10-303-204A-13
; Sequence 13, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10

```

```

; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 366
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-13

```

```

Query Match
Best Local Similarity 40.1%; Score 863.5; DB 12; Length 366;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWPPALPCD-----ERRCSPPLGALVPTAVCLCLFVGVSGN 56
Db 2 WNPATSEEPFNLTADLDWDASPGNDSLGLDELQLFPAPLLAGVATATCVAFVVGAGN 61
QY 57 VVTVMILIGRYDRMTTNNLYLGSMAVSDLLILLGLPDLVRLWRSRPWVFGPDLCLRLSLY 116
Db 62 LUTMLVWSRFRLEUTNNLYLSSMAFSDLLIFLCPDLVRLWQYRPNWPGDILLCKLQF 121
QY 117 VGEGCTYATLLHMTALSVERYLAICRPLRARLVTRRRVRLAVLWAVALLSAGPFLFL 176
Db 122 VSESCYATVLTALSVERYFAICFPRAKVVVTKRVLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVHEH-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFPLPCILSILYGLIGRELWSSRRPLRGPASGRGHR 294
Db 200 PTEFAVRSGLLTVMWVSSIFFFPLVFCVLTLYSLIGRLMRRRRGDVVGASLRDQNHK 259
QY 295 QTVRVLLVVLVAFIICWLPFHVGRIIYINTEDS---RMMYFSQYFNIVALQFLYLSASIN 351
Db 260 QIVKMLAVVFAFILLCPFHVGRIYLFKSPGSGLEIAQISQYCNLSVFLYLSAAN 319
QY 352 PILYNIISKYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTGYTETSAN 407
Db 320 PILYNIIMSKYRVAVFRLGLGFPFSQKSLSTLKDESSR-----AWTESSIN 365

```

RESULT 8

```

US-10-251-385-88
; Sequence 88, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 88
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-88

```

```

Query Match
Best Local Similarity 40.1%; Score 863.5; DB 15; Length 366;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

```

```
QY 5 WNGSDGPECA-----REPPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVGVSGN 56
DB 2 WNAATPESEPGFNLTLADLDWDASPGNDSGLDELQLLPAPLLAGVATATCAVLFVVGAGN 61
QY 57 VVTVMIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGPLLCRLSLY 116
DB 62 LITMLVVSFRRLTNNLYLSMAFSDDLIFLCMLPDLVRLWQYRPWFGDLCCKLQF 121
QY 117 VGEGETYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLALVAVALLSAGPFLF 176
DB 122 VSESCTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVFVIVAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVVPGCLNGTARIASSPLASSPPLWLSRAPPPSPGSPETABAAALFSRECR 236
DB 182 VGVGHE-----NGT-----DP-W-----DTNECR 199
QY 237 PS--PAQCALRVMLVMTTAYPPLPCLCLSLYGLIGRELSSRRPLRGPASGRERGRH 294
DB 200 PTEFAVRSGLTVMWWSIIFPFPVFCUTVLSLIGRLWRRRRGDVVGASLRQNHK 259
QY 295 QTVRVLLVVLAFIICWLPFHVGRIIYINTEDS---RMVFSQYFNIVALQLFYLSASIN 351
DB 260 QTVKMLAVVVFILCMLPFHVGRLYFSKSFEPGSLIAQISQYCNLVSFVLYLSAAN 319
QY 352 PLYNLISKYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSAN 407
DB 320 PLYNIMSKYRVAVFRLGLFEPFSQKLSLTKDESSR-----AWTESSIN 365

RESULT 9
US-10-303-204A-3
; Sequence 3, Application US/10303204A
; Publication NO. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigmer, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 353
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-303-204A-3

Query Match 40.0%; Score 863; DB 12; Length 353;
Best Local Similarity 46.5%; Pred. No. 4.1e-62;
Matches 186; Conservative 53; Mismatches 95; Indels 66; Gaps 9;

QY 19 WPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVGVSGNVTNNL 75
DB 8 WDAPPENDSLVEALLPFPPLLAGVATATCAVLFVVGAGNLLTMLVVSFRFMRITNL 67
QY 76 YLGSMAVSDLLILGLPFDLYRLWRSRPWFGPLLCRLSLYVGEGETYATLLHMTALSVE 135
DB 68 YLSSMAFSDLLIFLCMLPDLFRLWQYRPWNLGNLCKLQFVSESCTYATVLTITALSVE 127
```

```
QY 136 RYLAICRPLRARVLVTRRRVRLALVAVALLSAGPFLFVGVQDPGIVVVGNGTA 195
DB 128 RYFAICPPLRAKVVTKGRVKLVIVAVAFCSAGPIFVLVGVHED-----NGT- 177
QY 196 RIASSPLASSPPLWLSRAPPPSPGSPETABAAALFSRECRPS--PAQCALRVMLVMTT 253
DB 178 -----DPRD-----TNECRATEFAVRSGLLTVMWVSS 205
QY 254 AYFFLPFLCLSLYGLIGRELSSRRPLRGPASG---RERGHROTVRVLLVVLAFIIC 310
DB 206 VFFFLPVFCUTVLSLIGRLW---RRKGEAAVGSLLRDQNHKQTVKMLAVVVPFALIC 262
QY 311 WLPFHVGRIIY---INTEDSRMYPFSQYFNIVALQLFYLSASINPLYNLISKYRAAFA 367
DB 263 WLPFHVGRLYFSKLSLEPGSVEIAQISQYCNLVSFVLYLSAANPLYNIMSKYRVAVF 322
QY 368 KULLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSAN 407
DB 323 KULGPEPFSQKLSLTKDESSR-----AWTESSIN 352

RESULT 10
US-10-303-204A-16
; Sequence 16, Application US/10303204A
; Publication NO. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigmer, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: rattus norvegicus
US-10-303-204A-16

Query Match 40.0%; Score 861.5; DB 12; Length 364;
Best Local Similarity 48.3%; Pred. No. 5.7e-62;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGPEGARBP-----WPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVGVSG 55
DB 2 WNAAT--PSEPEPNNVTLDLWDASPGNDSLDPDELLPFPAPLLAGVATATCAVLFVVGISG 59
QY 56 NVVTVMIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGPLLCRLSL 115
DB 60 NLLTMLVVSFRRLTNNLYLSMAFSDLLIFLCMLPDLVRLWQYRPWFGDLCCKLQF 119
QY 116 YVGEGETYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLALVAVALLSAGPFLF 175
DB 120 FVSESCTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVIVAVAFCSAGPIFV 179
QY 176 LVGVQDPGIVVVGNGTARIASSPLASSPPLWLSRAPPPSPGSPETABAAALFSREC 235
DB 177 LVGVGHE-----NGT-----DP-W-----DTNECR 199
```

Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQLGALRVMLVWTTTAYFPFLPCLISLYGLIGRELWSSRPLRGPAAG--RE 290
Db 198 RATEFAVRSGLLTVWVSSVFFLPVFCVTVLSYLGRLKM--RR--RGAAGASLRD 253
QY 291 RGHRTQTVLLVVLAFIICWLPFHVGRRIIYINTEDS---RMYFSQYFNIVALQLFYLS 347
Db 254 QNHKQIVKMLAVVVFILCWLPHVGRYLFKSPFPGSLIAQISQYCNLVSVFLFYLS 313
QY 348 ASINPILYNISKYRAAAFKLL 370
Db 314 AAINPILYNIMSKYRVAVFKLL 336

RESULT 11
US-10-303-204A-8
; Sequence 8, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Atena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; FILE REFERENCE: 19589PCA
; CURRENT FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 361
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-8

Query Match 39.9%; Score 860; DB 12; Length 361;
Best Local Similarity 46.1%; Pred. No. 7.4e-62;
Matches 183; Conservative 57; Mismatches 97; Indels 60; Gaps 8;
QY 19 WPALPPCD---ERRCSPPLGALVPVTVAVCLCLFVVGSGNVVTVMLIGRYRDMRTTNL 75
Db 16 WDASPGNDSLGDELLQFPAPLAGVATCATVALFVVGAGNLLTMLVWSRFRRLRTTNL 75
QY 76 YLGSMVSDLLILGLPFDLYRLWRSPWVFGLLCRLSYVCEGCTYATLLHMTALSVE 135
Db 76 YLSSMAFSDLLIFLCMPDLVRLWQYRPMNFGDLCKLFQVSESCYATVLTITALSVE 135
QY 136 RYLAICRPLRARLVTRRRVRALIAVLWALLSAGPFLFVGVGQDPGISVVPGLNGTA 195
Db 136 RYFAICFPLRAKVTVTKRVLKLVFIWAVAFCSAGPIFVLVGVHE-----NGT- 185
QY 196 RIASSPLASPPPLWLSRAPPPSPGPETAFAALFSRCRPS--PAQLGALRVMLVWTT 253
Db 186 -----DP--W-----DTNECRPTEFAVRSGLLTVWVWVSS 213
QY 254 AYEFPLCLISLYGLIGRELWSSRPLRGPAAGRGHRTQTVLLVVLVLAFLICWLP 313
Db 214 IFPFLPVFCVTVLSYLGRLKRRRGDAVVGASLRDQNHKQIVKMLAVVVFILCWL 273
QY 314 FHVGRRIIYINTEDS---RMYFSQYFNIVALQLFYLSASINPILYNISKYRAAAFKLL 370

Db 274 FHVGRYLFKSPFPGSLIAQISQYCNLVSVFLFYLSAALNPLYNIMSKYRVAVFRL 333
QY 371 LARKSPRGRHRSRDTAGEVAGDTGGDTVGYTETSAN 407
Db 334 GFEPFSQRKSLTKDESSR-----AWTESSIN 360

RESULT 12
US-10-251-385-210
; Sequence 210, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 210
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-210

Query Match 39.8%; Score 857.5; DB 15; Length 366;
Best Local Similarity 44.5%; Pred. No. 1.2e-61;
Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;
QY 5 WNGSDGPEGA---REPPWPALPPCD---ERRCSPPLGALVPVTVAVCLCLFVVGSGN 56
Db 2 WNAATPSEEPFNLTADLDNDASPGNDSLGDELLQFPAPLAGVATCATVALFVVGAGN 61
QY 57 VVTVMILIGRYRDMRTTNLYLGSMVSDLLILGLPFDLYRLWRSPWVFGPPLLCRLSLY 116
Db 62 LLTMLVVSFRFRLRTTNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMNFGDLCKLFQ 121
QY 117 VEGGCTYATLLHMTALSVERYLAICRPLRARLVTRRRVRALIAVLWALLSAGPFL 176
Db 122 VSESCYATVLTITALSVERYFAICFPLRAKVTVTKRVLKLVFIWAVAFCSAGPIFVL 181
QY 177 VGVGQDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPETAFAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVWTTTAYFPFLPCLISLYGLIGRELWSSRPLRGPAAGRGHR 294
Db 200 PTEFAVRSGLLTVWVWVSSIFPFLPVFCVTVLSYLGRLKRRRGDAVVGASLRDQNHK 259
QY 295 QTVRVLLVVLVLAFLICWLPFHVGRRIIYINTEDS---RMYFSQYFNIVALQLFYLSIN 351
Db 260 QTKMLAVVVFVAFILCWLPHVGRYLFKSPFPGSLIAQISQYCNLVSVFLFYLSA 319
QY 352 PILYNLISKYRAAAFKLLARKSRPRGRHRSRDTAGEVAGDTGGDTVGYTETSAN 407
Db 320 PILYNIMSKYRVAVFRLIGFEPFSQRKSLTKDESSR-----AWTESSIN 365

RESULT 13
US-10-303-204A-2
; Sequence 2, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.


```
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: eus scrofa
US-10-303-204A-2

Query Match      36.0%; Score 775.5; DB 12; Length 302;
Best Local Similarity 46.6%; Pred. No. 4.4e-55;
Matches 166; Conservative 48; Mismatches 79; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSPWVFGPLLCRLSLYVGE 119
Db 1 MLVVSFRFRLRTTNNLYLSSMAFSDLLIFLCMPDLVRLQYRPMNGLNLLCKLQFVSE 60
QY 120 GCTYATLLHMTALSVERYLAIICPLARVLVTRRRVRLIALVAVALLSAGPFLVLGV 179
Db 61 SCTYATVLTITALSVERYFAICPLRAKVVTGRVKLVITVIAVAFCSAGPFLVLGV 120
QY 180 EODPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS- 238
Db 121 EHD-----NGT-----DP--W-----DTECRPTE 138
QY 239 -PAQALGALRVMLWVTYAYFPFLPCLSLYGLIGRELWSSRRPLRGPAASG-----RERGRH 294
Db 139 FAVRSGLLTVMVWSSVFFPFLVFCVLTVLYSLIGRKLWRRRRGDVAVGASLRDQNHK 195
QY 295 QTVRVLLVVVLAFLICWLPFHVGRIIY---INTEDSRMMYFQYFNIVALQFVLSASIN 351
Db 196 QTVKMLAVVVFALICWLPFHVGRIYFSKSPGSEVIAQISQYCNLVSPVLYLSAAIN 255
QY 352 PILYNIISKYRAAFAKLLARKSRPRGPHRSRDGTAGEVAGDTGGDTVGYTETSAN 407
Db 256 PILYNIISKYRVAVFKLLGFEPPFSQKLSLKDESSR-----AWTESSIN 301

RESULT 14
US-10-303-204A-7
; Sequence 7, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 271
; TYPE: PRT

; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 302
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-7

Query Match      35.9%; Score 774.5; DB 12; Length 302;
Best Local Similarity 46.5%; Pred. No. 5.3e-55;
Matches 164; Conservative 51; Mismatches 81; Indels 57; Gaps 7;

QY 60 VMLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSPWVFGPLLCRLSLYVGE 119
Db 1 MLVVSFRFRLRTTNNLYLSSMAFSDLLIFLCMPDLVRLQYRPMNFGDLLCKLQFVSE 60
QY 120 GCTYATLLHMTALSVERYLAIICPLARVLVTRRRVRLIALVAVALLSAGPFLVLGV 179
Db 61 SCTYATVLTITALSVERYFAICPLRAKVVTGRVKLVITVIAVAFCSAGPFLVLGV 120
QY 180 EODPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS- 238
Db 121 EHD-----NGT-----DP--W-----DTECRPTE 138
QY 239 -PAQALGALRVMLWVTYAYFPFLPCLSLYGLIGRELWSSRRPLRGPAASG-----RERGRH 297
Db 139 FAVRSGLLTVMVWSSVFFPFLVFCVLTVLYSLIGRKLWRRRRGDVAVGASLRDQNHK 198
QY 298 RVLLVVVLAFLICWLPFHVGRIIYINTEDS---RMMYFQYFNIVALQFVLSASINPIL 354
Db 199 KMLAVVVFALICWLPFHVGRIYFSKSPGSEVIAQISQYCNLVSPVLYLSAAINPIL 258
QY 355 YNLIISKYRAAFAKLLARKSRPRGPHRSRDGTAGEVAGDTGGDTVGYTETSAN 407
Db 259 YNIMSKYRVAVFKLLGFEPPFSQKLSLKDESSR-----AWTESSIN 301

RESULT 15
US-10-303-204A-12
; Sequence 12, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 271
; TYPE: PRT
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 26.3308 Seconds
(without alignments)
1504.757 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDGPGAREPPWP.....DTGGDTVGVTSTANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	23.2	424	2 JH0164	neurotensin recept
2	497	23.1	418	2 S29506	neurotensin recept
3	453	21.0	477	2 JC7313	capa receptor (CGI
4	437	20.3	418	2 A88013	protein K10B4.4 [i
5	401	18.6	416	2 S68822	neurotensin recept
6	375	17.4	378	2 T15816	hypothetical prote
7	359	16.7	363	2 I57940	somatostatin recep
8	357.5	16.6	418	2 A46226	somatostatin recep
9	344	16.0	363	2 I57955	somatostatin recep
10	344	16.0	364	2 JN0763	somatostatin recep
11	343.5	15.9	380	2 A55259	kappa opioid recep
12	339	15.7	352	2 JE0296	thyrotropin releas
13	339	15.7	380	2 S36143	kappa opioid recep
14	335	15.5	380	2 JC2338	kappa opioid recep
15	335	15.5	388	2 JN0605	somatostatin recep
16	334.5	15.5	367	2 I49022	kappa opioid recep
17	334.5	15.5	367	2 T24241	opioid receptor ho
18	334.5	15.5	367	2 I56520	G protein-coupled
19	334.5	15.5	428	2 S30508	probable G protein
20	331.5	15.4	428	2 A44021	somatostatin recep
21	330.5	15.3	370	2 S43087	orphan opioid rece
22	329	15.3	380	2 A48227	kappa opioid recep
23	328	15.2	380	2 JC2434	kappa opioid recep
24	327.5	15.2	384	2 A47249	brain-specific som
25	326.5	15.2	372	2 I38532	delta opioid recep
26	326	15.1	519	2 I37833	tachykinin recept
27	322	14.9	384	2 JC4629	somatostatin recep
28	321.5	14.9	372	2 S34592	delta opioid recep
29	321.5	14.9	398	2 JN0708	thyrotropin-releas

RESULT 1

JH0164

neurotensin receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000

C;Accession: JH0164

R;Tanaka, K.; Masu, M.; Nakanishi, S.

Neuron 4, 847-854, 1990

A;Title: Structure and functional expression of the cloned rat neurotensin receptor.

A;Reference number: JH0164; MUID:90297956; PMID:1694443

A;Accession: JH0164

A;Molecule type: mRNA

A;Residues: 1-424 <TAN>

C;Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor. This

cter (neuromodulator) in the brain and as a hormone) cellular mediator in peripheral tiss

C;Superfamily: vertebrate rhodopsin

C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F;65-87/Domain: transmembrane #status predicted <TM1>

F;97-121/Domain: transmembrane #status predicted <TM2>

F;144-165/Domain: transmembrane #status predicted <TM3>

F;189-210/Domain: transmembrane #status predicted <TM4>

F;236-260/Domain: transmembrane #status predicted <TM5>

F;309-330/Domain: transmembrane #status predicted <TM6>

F;348-372/Domain: transmembrane #status predicted <TM7>

F;4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 23.2%; Score 501; DB 2; Length 424;

Best Local Similarity 32.9%; Pred. No. 2.4e-32;

Matches 125; Conservative 59; Mismatches 110; Indels 86; Gaps 11;

QY	39	VPTAVTCLFLVVGVSNGVTVMLIGR---	YRDMRTTNLYLGSMAVSDLLI-LLGLPFD 94
DB	65	VLVTATYALFVVGTVGTVGNSTFTLARKKSLSQSLQSTVHYHLGLSLALSLLILLAMPVE 124	
QY	95	LVR-LWRSPWFGPLCLRLSLVVGCGCTVATLLHMTALSVRYLAICRLRLRVLTTR 153	
DB	125	LYNFIWHPHPAFGDAGCGYFLRDACYATATNALNVASLSVERYLAICHPFKATLMRSR 184	
QY	154	RVRALIAVLWAVALLSAGPFLPLVGVQDPGISVWPGLNGTARIASSPLASSPPLMSRA 213	
DB	185	RTKKFISATWLASALLAIPMLFTMGLQNRSRGDGTGPG-----GLVCTPIVDT----- 231	
QY	214	PPSPSPSGPETABAAALFSRECRPSPAQLGALRVMLVWVTAYTFFL-PPFLCLSLYLGILGR 272	
DB	232	-----ATVKVVIQVNTFMSPLFPMLVISILNTVIAN 262	
QY	273	ELWSSRRRLRGPAAAGR-----ERGHQTVR---VLLVVLAFII 309	
DB	263	KLTVM---VHQAAEQGRVCTVGTGTHNGLEHSTFNWTEPGRVQALRGVLRVAVTAFV 319	
QY	310	CWLPPFHVGRIIYINTEDSR-----MMYFSQYFNIVALQLPFLYSASINPILYNLISKYRAA 365	

delta opioid recep
galanin receptor 2
thyrotropin-relea
thyrotropin recep
angiotensin recep
thyrotropin-releas
neurokinin 2 recep
somatostatin recep
delta opioid recep
neurokinin 1 recep
somatostatin recep
neurokinin 1 recep
neurokinin 2 recep
somatostatin recep
opioid receptor mu

ALIGNMENTS

```
Db 320 CWLPYHVRRLMFCYISDEQWTTFLDFYHYFYMLTNLFYVSSAINPILNLYSANFRQV 379
QY 366 AFKLLLA-----RKSRP 377
Db 380 FLSTLACLCPGWRHRRKKRP 399

RESULT 2
S29506
neurotensin receptor - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C:Accession: S29506
R;Vita, N.; Laurent, P.; Lefort, S.; Chalon, P.; Dumont, X.; Kaghad, M.; Gully, D.; le H
PEBS Lett. 317, 139-142, 1993
A;Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
A;Reference number: S29506; MUID:93154505; PMID:8361365
A;Accession: S29506
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-418 <VT>
A;Cross-references: EMBL:X70070; NID:g35020; PIDN:CAA49675.1; PID:g35021
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 23.1%; Score 497; DB 2; Length 418;
Best Local Similarity 33.1%; Pred. No. 5e-32;
Matches 126; Conservative 62; Mismatches 117; Indels 76; Gaps 13;

QY 39 VPVTAVCLFVVGSGNVVTVMLIGR---YRDMRTTNNLYLGSMAVSDLL-ILLGLPFD 94
Db 64 VLVTVAVLALFVVGTVGNTVTAFTLARKKSLSQSLQSTVHYHLGSLALSDDLTLAMPVE 123
QY 95 LYR-LWRSRPNVGPPLLCRLSLVGEQCTYATLLHMTALSVERYLAICRPLARVLVTRR 153
Db 124 LYNFIVHHPWAFGACRGYFLRDCTATALNVASLSVERYLAICHPFKAKTLMRS 183
QY 154 RVRLALIAVLWALLSAGPLFLVGVQDPGISVVPGLNGTARIASSPLASSPPLWLSRA 213
Db 184 RTKKFSAIWLASALLTVMLFTWG-EQN----- 211
QY 214 PPPSPSGPTABAAALFSRECRPSAQLGALRVMLVWTT-AYFFLPFLCLSLYGLIGR 272
Db 212 -----RSADGQAGGLVCTPT-IHTATVKVIQNTFMGSFIFPMVVISVLNTIIAN 261
QY 273 ELWSSRPL--RGPAA-----GREGHQRTVRLVVLVLAFLICWLPFHV 316
Db 262 KLTVMVRQAABEQVCTVGEHSTFSAIEPGRVQALRHGVRVLRVAVIAFVVCWLPFHV 321
QY 317 GRIIYNTDSR--MMY-PSQYFNIVALQLFYLASINPILNLYSKYR-----AA 365
Db 322 RRLMFCYISDEQWTFDYHYFYVMYTNALFVSSSTINPILNLYSANFRHIFLATLAC 381
QY 366 AFKLLARKSRPRGRHRSRT 386
Db 382 LCPWRRRRKKRP-AFSRKADS 401

RESULT 3
JC7913
capa receptor (CG14575) - fruit fly (Drosophila sp.)
C:Species: Drosophila sp.
C>Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 31-Mar-2003
C:Accession: JC7913
R;Iversen, A.; Cazzamali, G.; Williamson, M.; Hauser, F., and Grimmelikhuijzen, C.J.P.
Biochem. Biophys. Res. Commun. 299, 628-633, 2002
A;Title: Molecular cloning and functional expression of a Drosophila receptor for the ne
A;Reference number: JC7913; MUID:22347021; PMID:12459185
A;Accession: JC7913
A;Molecule type: mRNA
A;Residues: 1-477 <IVE>
A;Cross-references: GB:AF505865
C;Comment: This receptor is a G-protein-coupled receptor stimulates renal (Maipighi)
```

```
C;Genetics:
A;Introns: 70/3; 121/2; 126/1; 218/2; 276/2; 301/2; 349/1; 402/3; 436/3

Query Match 21.0%; Score 453; DB 2; Length 477;
Best Local Similarity 29.7%; Pred. No. 1.8e-28;
Matches 117; Conservative 68; Mismatches 135; Indels 74; Gaps 11;

QY 8 SDGPEGAREPPWPPALPPCDERRCSP-----FPLGALVPVTAVCCLFVVGVS 55
Db 30 SDPSHGGEEDY-----ACGTFNCSPKEFVAVLGPQTLPLYKAVLITIFGGIFITGVG 85
QY 56 NVVTVMILIGRYDRMRTTNNLYLGSMAVSDLLILL-GLPPDLYLRKSRPWVFPGLLCRLS 114
Db 86 NLLVCIVIRHSAMHTATNYFLSLAVSDLLYLLFGLTEVFLYMHQYDPDFGMPFKIR 145
QY 115 LYVGECTYATLLHMTALSVERVLAICRPLARVLTTRVRVRLIAVLWALLSAGPFL 174
Db 146 AFISEACTYVSVFTVAFSMERFLAICHLPHLHYAMVGFKRAIRIITALTMTVSVISAI 204
QY 175 FLVGVQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAFAALFSRE 234
Db 205 -----GLLSDIQVLYNPLDHS-----RIEESAF----- 227
QY 235 CRPSAQLGALRVMLVWTTAYFPFLCFLSLYGLIGRELWSSRRPLRG--PAASGRERG 292
Db 228 CSMSPKIVNEIPVFEVSFCITFFVPMILILLYGRMGAKIRSRTNOKLGVOQGTNNRETR 287
QY 293 HRQ-----TVRLVVVLAFLICWLPFHVGRITTYINTEDSRMMYFSQYFNI-----VAL 341
Db 288 NSQMRKXTVIRMLAAVVIITFCWFFHLQRLIFYAKN-----MDNYLDINEALFSIAG 342
QY 342 QLFLYSASINPILNLYSKYRAAFAKLLLARKS 375
Db 343 FAYVSVCTVNPVYSVMSRRYR-VAPRELLCGKA 375

RESULT 4
A88013
protein K10B4.4 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: A88013
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: A88013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-418 <STO>
A;Cross-references: GB:chr_II; PIDN:AB71009.1; PID:g2429475; GSPDB:GN00020; CESP:K10B4.4
A;Note: similar to family I of G-protein coupled receptors
C;Genetics:
A;Gene: K10B4.4
A;Map position: 2

Query Match 20.3%; Score 437; DB 2; Length 418;
Best Local Similarity 30.1%; Pred. No. 2.9e-27;
Matches 109; Conservative 78; Mismatches 137; Indels 38; Gaps 10;

QY 29 RCSPPFLGALVPVTAVCCLFVVGSGNVVTVMLIGRYDRMRTTNNLYLGSMAVSDLL-I 87
Db 24 RCQ--SAGIVIPVTIYIGTIFLLGLFGNICTCIVIAANKSMHNPNTNYLFLSAVSDIAL 81
QY 88 LLGLPPDLXR-LWRSRPNVGPPLLCRLSLYVGEQCTYATLLHMTALSVERYLAICRPLA 146
Db 82 ILGLPMFQSLDYSYPRFSEGCICARAFLEFTSYASIMIICCFSEFWLAIChPLRS 141
QY 147 RVLVTRRRVRALIAVLWALLSAGPFLFLVGVQDP-----GISVVPGLNGTARIASSPLA 203
Db 142 KIPSTLWRANVLIIIAWTISFVCALDIAFIVQINKLPLPEDAKYQPWTKVSPFAVGVLN 201
```


A;Reference number: I57949; MUID:94088493; PMID:8264565
A;Accession: I57949
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 341-363 <OCA2>
A;Cross-references: GB:867370; NID:G455947; PIDN:AA29371.1; PID:G455948
A;Experimental source: pituitary
R;Penetta, R.; Greenwood, M.; Patel, Y.C.
submitted to the EMBL Data Library, August 1993
A;Description: Correction of the nucleotide and amino acid sequence of the rat somatostatin receptor.
A;Reference number: S3244
A;Accession: S3244
A;Molecule type: mRNA
A;Residues: 309-363 <PEN>
A;Cross-references: EMBL:X74828; NID:G433911; PIDN:CAA52825.1; PID:G433912
C;Genetics:
A;Gene: SSTP5
C;Superfamily: vertebrate rhodopsin

Query Match 16.7%; Score 359; DB 2; Length 363;
Best Local Similarity 27.3%; Pred. No. 4.1e-21;
Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;

QY 5 WNGSDGEGAREPPWALPCDERRCSPFLGA---LVPVAVCLCLFVVGVSGNVVVM 61
DB 12 WNASAAAGSNHN--WSLVG-----SASPMGARAVLPVLYLLVC--TVGLSGNTLVY 60

QY 62 LIGRYDMRTTNLYGSMVAVDLLILGLPFDLYLRWRSPWVFGPLLCRLSLYVGE 121
DB 61 VTLRHAKMTVTNVIILNADVLFMLGLPFLATONAVVSYPPFGSFLCRLVNTLDGIN 120

QY 122 TYATLHMTALSVERYLAICRPLARVLVTRRRVRLIALIWAVALLSAGPFLFLVGVEQ 181
DB 121 QFTSIFCLWMSVDRLAVVHPTRASRRPRVAKASAAVWFSLLMSLPLLVFADVQE 180

QY 182 DPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFRCRPSAQ 241
DB 181 G-----WGTCNLS-----W-----PEPVG 194

QY 242 LGALRVMLWVTAYFFLPFLCLSLYGLI-----GRELWSSRRPLRGPAASGRERCHR 294
DB 195 LWGAFTYTSVLGFFGFLVLCYLLIVVKVKAAGMRVGSRR-----RRSE 244

QY 295 QTVRVLLAVVLAFTICMLPHVGRRIIYN--TEDSRMMYFSQYFNIVALQLFYLSASINPI 353
DB 245 KVTWVWVWVVLVFGVCLPFFIVNVLATLPEEPTSGAGLYFVVVLS--YANSCANPL 302

QY 354 LYNLSKKYRAAFLKLLARKSRPRGPHRSRDTAGEVAGDTGG 396
DB 303 LYGLFSDNFRQSPRKVLCLR---RGYGMEDADAIEPRPDKSG 341

RESULT 8
A46226
somatostatin receptor 3 - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C;Accession: A46226; S32501
R;Yamada, Y.; Reisine, T.; Law, S.F.; Ihara, Y.; Kubota, A.; Kagimoto, S.; Seino, M.; Seino, M.; Endocrinol. 6, 2136-2142, 1992
A;Title: Somatostatin receptors, an expanding gene family: cloning and functional characterization.
A;Reference number: A46226; MUID:93149123; PMID:1337145
A;Accession: A46226
A;Molecule type: DNA
A;Residues: 1-418 <YAM>
A;Cross-references: GB:M96738; NID:G338498; PIDN:AAA60592.1; PID:G338499
A;Note: sequence extracted from NCBI backbone (NCBIN:123685, NCBIPI:123690)
R;Corness, J.D.; Demchyshyn, L.L.; Seeman, P.; van Tol, H.H.M.; Srikant, C.B.; Kent, G.; FEBS Lett. 321, 279-284, 1993
A;Title: A human somatostatin receptor. (SSTR3), located on chromosome 22, displays preference for the [D-Phe1]somatostatin.
A;Reference number: S32501; MUID:93238970; PMID:8097479
A;Accession: S32501
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-418 <COR>
C;Genetics:
A;Gene: GDB:SSTR3
A;Cross-references: GDB:134187; OMIM:182453
A;Map position: 22q13.1-22q13.1
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F;44-70/Domain: transmembrane #status predicted <TM1>
F;81-106/Domain: transmembrane #status predicted <TM2>
F;118-139/Domain: transmembrane #status predicted <TM3>
F;159-181/Domain: transmembrane #status predicted <TM4>
F;203-233/Domain: transmembrane #status predicted <TM5>
F;255-282/Domain: transmembrane #status predicted <TM6>
F;289-316/Domain: transmembrane #status predicted <TM7>
F;17,30/Binding site: carboxylate (Asn) (covalent) #status predicted
F;116-191/Disulfide bonds: #status predicted
F;151,251,317,332/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;256/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predicted
F;412/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 16.6%; Score 357.5; DB 2; Length 418;
Best Local Similarity 26.9%; Pred. No. 6.2e-21;
Matches 101; Conservative 63; Mismatches 146; Indels 65; Gaps 10;

QY 11 PEGAREPPWALPCDERRCSPFL-----GALVPTAVCLCLFVVGVSGNVVTVMLIGR 65
DB 15 PENA--SSAWPPDPTLGNVAGSPAGLAVSGVLIPLVLYVC--VVGLLGNLSLVIVVLR 71

QY 66 YRDMRTTNLYGSMVAVDLLILGLPFDLYLRWRSPWVFGPLLCRLSLYVGE 125
DB 72 HTASPSVTNVIILNADVLFMLGLPFLAAQNALSYWFFGSLMCLWVAVDGINQPTS 130

QY 126 LHMTALSVERYLAICRPLARVLVTRRRVRLIALIWAVALLSAGPFLFLVGVE 185
DB 131 IFCLTVMVSDRLAVVHPTRASRRPRVAKASAAVWFSLLMSLPLLVFADVQE 184

QY 186 SVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFRCRPSAQ 242
DB 185 -----PRGMS-----CHMQWPEPAAA 201

QY 243 GALRVMLWVTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRERHRTVRVLLV 302
DB 202 WRAGFIYTAALGFPGFLVLCYLLIVVKVRSAGRWAPSCQRRRRSERRVTRMVA 261

QY 303 VLAFIICWLPFHVGRRIIYNTE--DSRMMYFSQYFNIVALQLFYLSASINPIYLSISK 361
DB 262 VVALFVLCWMPFYVLNIVNVCPLPEEPAPFGLYFLVALP--YANSCANPILYGLFSYR 319

QY 362 YRAAFAKLLARKSR 376
DB 320 FK-QGFRVLLRPSR 333

RESULT 9
I57955
somatostatin receptor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I57955
R;Panetta, R.; Greenwood, M.T.; Warszynska, A.; Demchyshyn, L.L.; Day, R.; Niznik, H.B.; Mol. Pharmacol. 45, 417-427, 1994
A;Title: Molecular cloning, functional characterization, and chromosomal localization of the human somatostatin receptor.
A;Reference number: I57955; MUID:94195267; PMID:7908405
A;Accession: I57955
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-363 <RES>
A;Cross-references: GB:L14865; NID:G431094; PIDN:AAA20828.1; PID:G431095
C;Genetics:
A;Gene: GDB:SST

A;Cross-references: GDB:119604; OMIM:182450
A;Map position: 3q28-3q28
C;Superfamily: vertebrate rhodopsin

Query Match 16.0%; Score 344; DB 2; Length 363;
Best Local Similarity 27.7%; Pred. No. 6.4e-20;
Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;
QY 14 AREPPWPALPP-----CDERR-CSPPP-LGA-----LVPVTAVCLCLFVVGSGNVVTMVL 62
DB 7 ASTPSWNASPGASGGDNRTLVGPAPSGARAVLPVLYLLVC--AAGLGGNTLIYIV 64
QY 63 IGRYDRMTTNYLYGSMVSDLLILGLPFDLYRLWRSRPWPGPCLRLSLYVSGECT 122
DB 65 VLRFAMKTKVTNIYILNLAADVLYMLGLPP-LATQNAASFMPGPGVLCRLVMTLDGVNQ 123
QY 123 YATLLHMTALSVRYLAICPLRARVLVTRRRVRLALAVLWAVALLSAGPFLFLVGEQD 182
DB 124 FTSVFCUTVMSVDYLAIVHPLSSARWRPRVAKLASAAWVLSCLMSLPLLVFADVOE- 182
QY 183 PGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETABAAALFSRECRPSPAQL 242
DB 183 -----GGTCNAS-----W-----PEPVL 196
QY 243 GALTVMVMTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPASG-----BERGHRQT 296
DB 197 WGAFFIITYAVLGFAPFLVICLCYLLIVKV-----RAAGVRVGCVRSSRKV 246
QY 297 VRVLVVLVLAFLICWLPFHVGRIIYI-----NTEDSRMYPFQYFNIVALQLFVLSASIN 351
DB 247 TRMLVVLVLFAGCWLPPFTTNIWNLAVLPQSPASAGLYP---FVVI---LSYNSCAN 300
QY 352 PILYNLISKYRAAFAKLLARK 374
DB 301 PVLVGLFSDNFRQFQVCLRK 323

RESULT 10
JN0763
somatostatin receptor 5 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
C;Accession: JN0763
R;Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A;Title: Cloning, functional expression and pharmacological characterization of a fourth
A;Reference number: JN0762; MUID:93384611; PMID:8373420
A;Accession: JN0763
A;Molecule type: DNA
A;Cross-references: DDBJ:D16827; NID:9487683; PIDN:BAA04107.1; PID:9487684
C;Comment: This protein is a member of somatostatin receptor family.
C;Genetics:
A;Gene: GDB:SSTR5
A;Cross-references: GDB:138452; OMIM:182455
A;Map position: 16p13.3-16p13.3
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; phosphoprotein; thiol
F;40-66/Domain: transmembrane #status predicted <TM1>
F;77-102/Domain: transmembrane #status predicted <TM2>
F;114-135/Domain: transmembrane #status predicted <TM3>
F;155-177/Domain: transmembrane #status predicted <TM4>
F;196-228/Domain: transmembrane #status predicted <TM5>
F;246-273/Domain: transmembrane #status predicted <TM6>
F;280-307/Domain: transmembrane #status predicted <TM7>
F;13,26,187/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;112-186/Diulfide bonds: #status predicted
F;242-325/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pr
F;247/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predic
F;320/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 16.0%; Score 344; DB 2; Length 364;

Best Local Similarity 27.7%; Pred. No. 6.4e-20;
Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;
QY 14 AREPPWPALPP-----CDERR-CSPPP-LGA-----LVPVTAVCLCLFVVGSGNVVTMVL 62
DB 7 ASTPSWNASPGASGGDNRTLVGPAPSGARAVLPVLYLLVC--AAGLGGNTLIYIV 64
QY 63 IGRYDRMTTNYLYGSMVSDLLILGLPFDLYRLWRSRPWPGPCLRLSLYVSGECT 122
DB 65 VLRFAMKTKVTNIYILNLAADVLYMLGLPP-LATQNAASFMPGPGVLCRLVMTLDGVNQ 123
QY 123 YATLLHMTALSVRYLAICPLRARVLVTRRRVRLALAVLWAVALLSAGPFLFLVGEQD 182
DB 124 FTSVFCUTVMSVDYLAIVHPLSSARWRPRVAKLASAAWVLSCLMSLPLLVFADVOE- 182
QY 183 PGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETABAAALFSRECRPSPAQL 242
DB 183 -----GGTCNAS-----W-----PEPVL 196
QY 243 GALTVMVMTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPASG-----BERGHRQT 296
DB 197 WGAFFIITYAVLGFAPFLVICLCYLLIVKV-----RAAGVRVGCVRSSRKV 246
QY 297 VRVLVVLVLAFLICWLPFHVGRIIYI-----NTEDSRMYPFQYFNIVALQLFVLSASIN 351
DB 247 TRMLVVLVLFAGCWLPPFTTNIWNLAVLPQSPASAGLYP---FVVI---LSYNSCAN 300
QY 352 PILYNLISKYRAAFAKLLARK 374
DB 301 PVLVGLFSDNFRQFQVCLRK 323

RESULT 11
A55259
kappa opioid receptor - guinea pig
N;Alternate names: dynorphin receptor
C;Species: Cavia porcellus (guinea pig)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 24-Nov-1999
C;Accession: A55259
R;Xie, G.; Meng, F.; Mansour, A.; Thompson, R.C.; Hoversten, M.T.; Goldstein, A.; Watson, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 3779-3783, 1994
A;Title: Primary structure and functional expression of a guinea pig kappa opioid (dynor
A;Reference number: A55259; MUID:94224825; PMID:8170987
A;Accession: A55259
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-380 <XIE>
A;Cross-references: GB:U04092; NID:9476106; PIDN:AAA67171.1; PID:9476107
C;Superfamily: vertebrate rhodopsin
C;Keywords: transmembrane protein

Query Match 15.9%; Score 343.5; DB 2; Length 380;
Best Local Similarity 25.1%; Pred. No. 7.3e-20;
Matches 102; Conservative 74; Mismatches 147; Indels 83; Gaps 12;
QY 2 GSPW-----NGSDGPGAREPP---WPALPPCDERRCSPPLGALVPVAVCLCLF 49
DB 26 GSAWLPQWAEPCDNGSAGPQDELEPAHISPAIP-----VIIITAVYSVP 70
QY 50 VVGSGNVVTMVLIGRYDRMTTNYLYGSMVSDLLILGLPFDLYRLWRSRPWVFGPL 109
DB 71 VUGLVGNSLVMPFVIIRYTKMTATNIYIFNLADALAVTTTTPFQ-STVYLMSNSWPFQDV 129
QY 110 LCRSLYVVGSGCTYATLLHMTALSVRYLAICPLRARVLVTRRRVRLALAVLWAVALLS 169
DB 130 LCRIVISIDYNNFTSIFLTMTMSVDRIYAVCHPVKALDPRFLKAKIINICWLLSSSV 189
QY 170 AGPFLVGVGEQDQPGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETABAAA 229
DB 190 GISAILGGTKVREDVDIIE-----CSLQPDDDYSWD----- 223
QY 230 LFSRECRPSPAQLGALTVMVMTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPASGR 289

Db 224 LFWKIC-----VFVFAFVPLVLIIVCYTLMILRL-KSVALLSG--SREK 265

QY 290 ERGHRQTRVLLVVLAFIICWLPFHVGRIIYI--NTEDSRMMVFSQYFNIVALQLFYLS 347

Db 266 DNLRLRITRLVVLVAVFICWPIFIHIFILVEALGSTSHSTAALSSYF---CIALGYTN 322

QY 348 ASINPILYNLISKYRAA-----AFKLLARKSRPRGPHRSRDTA 387

Db 323 SSLNPILYAFDENFKRCDPFCFPPIKMRMERQSTSRVNTVQDPA 368

RESULT 12

JE0296

thyrotropin releasing hormone receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000

C;Accession: JE0296

R;Ishihara, H.; Nakamura, T.; Itoh, J.; Iwase, H.; Kanatani, A.; Borkowski, J.; Ihara, M. Biochem. Biophys. Res. Commun. 250, 68-71, 1998

A;Title: Cloning and characterization of a new subtype of thyrotropin-releasing hormone A;Reference number: JE0296; MUID:98407892; PMID:9735333

A;Accession: JE0296

A;Molecule type: mRNA

A;Residues: 1-352 <ITA>

A;Cross-references: DDBJ:AB015645; NID:G3660553; PIDN:BAA33437.1; PID:G3660554

C;Superfamily: adenosine receptor A1

F;26-48/Domain: transmembrane #status predicted <TM1>

F;58-80/Domain: transmembrane #status predicted <TM2>

F;97-118/Domain: transmembrane #status predicted <TM3>

F;142-165/Domain: transmembrane #status predicted <TM4>

F;188-209/Domain: transmembrane #status predicted <TM5>

F;252-273/Domain: transmembrane #status predicted <TM6>

F;282-304/Domain: transmembrane #status predicted <TM7>

Query Match 15.7%; Score 339; DB 2; Length 352;

Best Local Similarity 28.8%; Pred. No. 1.5e-19;

Matches 108; Conservative 52; Mismatches 109; Indels 106; Gaps 13;

QY 41 VTAVCLCLFY--VGVSNGVTVMLIGRYDMRTTNNLYGSMVSDLLILL--GLPPDLY 96

Db 23 VVSFLVLLVCTGIVGNAMVILVLTSDMHTPTNCYLVSLADALLVLAAGLPNVSD 82

QY 97 RLWRSRPWFVGLCLRLSLXVGGCTYATLL-----HMTLSVERYLAICRPLRARVL 149

Db 83 SL--VGHWVYGRAGCL-----GITVQYVLGINVSSFSILFTVERYIAICPLRAQTV 133

QY 150 VTRRRVRLTAVLWVALLSAGPFLVLGV-----EQDPSVVPGLNGTARIASPLA 203

Db 134 CTVARAKRIIAGVGTSLYCLLWFFLVLDLNVDRNQRLGCKYKVRGL-----181

QY 204 SSPPLMLSRAPPSPSPGPETAEEALFSRECRPSAQLGALRVMLVWTTTAYFFLPFLCL 263

Db 182 -----YLPYLLDFPFFIGPLVLT 201

QY 264 SILYGLIGR-----ELWSSRRPLRG--PAASGR-----ERGHQTRVRLVVLVLAFL 308

Db 202 LVLVGLIGRLIFQSPLSQEAQXQERQPHGQSEAPAGNCSRAKSRKQATRLAVVLLFA 261

QY 309 ICWLPFHVGRIIYI--NTEDSRMMVFSQYFNIVALQLF-----YLSASINPILYNLISKYR 363

Db 262 VLATPYRT--LVLLNS-----FVAQPLDPWLLFCRTCVTNSAVNPVYSLMSQKR 313

QY 364 AAAPKLLARKSRPR 378

Db 314 AAFKLWCRAAGPQ 328

RESULT 13

S36143

kappa opioid receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 10-Dec-1993 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000

C;Accession: S36143; S38825; S36102; S39015; A48789

R;Nishi, M.; Takeshima, H.; Fukuda, K.; Kato, S.; Mori, K. FEBS Lett. 330, 77-80, 1993

A;Title: cDNA cloning and pharmacological characterization of an opioid receptor with high affinity for morphine; A;Reference number: S36143; MUID:93380575; PMID:8396539

A;Accession: S36143

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-380 <NIS>

A;Cross-references: GB:D16534; NID:G409390; PIDN:BAA03971.1; PID:G415310

R;Chen, Y.; Mestek, A.; Liu, J.; Yu, L. Biochem. J. 295, 625-628, 1993

A;Title: Molecular cloning of a rat kappa opioid receptor reveals sequence similarities to the human kappa opioid receptor; A;Reference number: S38825; MUID:94059008; PMID:8240267

A;Accession: S38825

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-380 <CHE>

A;Cross-references: GB:I22001; NID:G409236; PIDN:AAA1495.1; PID:G409237

R;Minami, M.; Toya, T.; Katao, Y.; Maekawa, K.; Nakamura, S.; Onogi, T.; Kaneko, S.; Satoh, T. FEBS Lett. 329, 291-295, 1993

A;Title: Cloning and expression of a cDNA for the rat kappa-opioid receptor. A;Reference number: S36102; MUID:93374033; PMID:8103466

A;Accession: S36102

A;Molecule type: mRNA

A;Residues: 1-41, 'L', '43-380 <MIN>

A;Cross-references: GB:D16829; NID:G404115; PIDN:BAA04109.1; PID:G404116

R;Li, S.; Zhu, J.; Chen, C.; Chen, Y.W.; Deriel, J.K.; Ashby, B.; Liu-Chen, L.Y. Biochem. J. 295, 629-633, 1993

A;Title: Molecular cloning and expression of a rat kappa opioid receptor. A;Reference number: S39015; MUID:94059009; PMID:8240268

A;Accession: S39015

A;Molecule type: mRNA

A;Residues: 1-344, 'Y', '346-380 <LIS>

A;Cross-references: GB:I22536; NID:G425188; PIDN:AAA1496.1; PID:G425189

R;Meng, F.; Xie, G. Proc. Natl. Acad. Sci. U.S.A. 90, 9954-9958, 1993

A;Title: Cloning and pharmacological characterization of a rat kappa opioid receptor. A;Reference number: A48789; MUID:94052210; PMID:8234341

A;Accession: A48789

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-380 <RES>

A;Cross-references: EMBL:U00442; NID:G403486; PIDN:AAA18261.1; PID:G403487

C;Superfamily: vertebrate rhodopsin

C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 15.7%; Score 339; DB 2; Length 380;

Best Local Similarity 25.5%; Pred. No. 1.7e-19;

Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;

QY 6 NGSDDGPEGAREPP--WPALPPCDERRCSPFFPLGALVPVTVAVCLCLFVVGVSNGVTVML 62

Db 39 NGSVGSEDOOLEPAHISPAIP-----VIITAVSVVVFVGLVNSLVWFV 83

QY 63 IGRYDMRTTNNLYGSMVSDLLILLGLPFDLYRLRSRPPWVFGPPLCLRLSLYVGGCT 122

Db 84 IIRYTKTKATNIYIENLADALVTTTHFPQ--SAVVLNMSWFPGLVCLKIVISIDYNN 142

QY 123 YATLLHMTALSVRYLAICRPLRARVLVTRRRVRLTAVLWVALLSAGPFLVLVGEQD 182

Db 143 FTSIFTLTWSVDRYTAVCHPVKALDFRTPKAKIINICIWLLASSVGISAIVLGGTKVR 202

QY 183 PGISVVPGLNGTARIASSPLASSPPLWLSRAPPSPPGPETAEEALFSRECRPSAQL 242

Db 203 EDVDVIE-----CSLOFPDDEYSWMD-----LFMKIC-----229

QY 243 GALTVMVLTWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGHRQTRVRLV 302

Db 230 -----VFVFAFVPLVLIIVCYTLMILRL-KSVRLLSG--SREKDLNRITKLVLV 278

QY 303 VLAFIICWLPFHVGRIIYI--NTEDSRMMVFSQYFNIVALQLFYLSASINPILYNLISK 360

Db 279 VVAVFIICWPIFIHIFILVEALGSTSHSTAALSSYF---CIALGYTNSSLPVLYAFLDE 335

Qy 123 YATLLHWTALSVRYLACRPLRARVLVRRRRVALIAVLAVALLSAGPFLFVGVEOD 187
D :
Db 143 FTSITLTMMSDVRYYIVACHPVKALDFRTPLUKAKIINICIIWLLSSVGISAIVLGCTKVR 202
D :
Qy 183 PGI SVVPGLNGTARIASSPLASSPPLWLSRAPPPSPSGPETABAAALPSRECRPSPAQL 242
D :
Db 203 EDVDVIE-----CSLQPDDDDYSNWD-----LPMKIC----- 229
Qy 243 GALLRVMLWVTYAYFFLPFLCLSLYLIGILGWLSRRRLRGPAASGRHRHQTRVRLVLY 302
D :
Db 230 -----VIPAPVPIVLIIVCYTLMILRL-KSVRLLSG--SREKORNLRRIIRTLVLV 278
Qy 303 VVLAFTICWLPHFHVGRYYI--NTEDSRMYFSQFNVAIQLFYLSASINPILYNLISK 360
D :
Db 279 VVAVFVVCWTPFIHFILFIVEALGSTSHSTAALSSYYF---CIALGYTNSLNIPLYAFLOE 335
Qy 361 KYRAA-----AFKLLARKSRPRGFHSRDTA 387
D :
Db 336 NFKRCFRDPFCPLKKRMEROSTSRVRTNVQDPA 368

RESULT 15
JN0605
somatostatin receptor 4 - human
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 20-Jun-2000
C/Acession: JN0605; JN0762; A47457
R/Xu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.
Biochem. Biophys. Res. Commun. 193, 648-652, 1993
A>Title: Molecular cloning and sequencing of a human somatostatin receptor, hS
A/Reference number: JN0605; MUID:93290656; PMID:8512564
A/Acession: JN0605
A/Molecule type: DNA
A/Residues: 1-388 <XY>
A/Cross-references: GB:I14856; NID:92242499; PIDN:AAA36623.1; PID:g292500
R/Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ih
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A>Title: Cloning, functional expression and pharmacological characterization o
A/Reference number: JN0762; MUID:93384611; PMID:8373420
A/Acession: JN0762
A/Molecule type: DNA
A/Residues: 1-388 <YAM>
A/Cross-references: GB:D16826; NID:9693907; PIDN:BAA04106.1; PID:g693908
R/Rohrer, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, P.; Schule, R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
A>Title: Cloning and Characterization of a fourth human somatostatin receptor.
A/Reference number: A47457; MUID:93248256; PMID:8483934
A/Acession: A47457
A/Molecule type: DNA
A/Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>
A/Cross-references: GB:L07833; NID:9307429; PIDN:AAA60565.1; PID:g307430
A/Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBI:P:130856)
C/Comment: This protein mediates the diverse actions of the tetradecaptide som
C/Genetics:
A/Gene: GDB:SSTR4
A/Cross-references: GDB:202662; OMIM:182454
A/Map position: 20p11.2-20p11.2
A/Introns: #status absent
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipopr
P/F47-73/Domain: transmembrane #status predicted <TM1>
F/F84-109/Domain: transmembrane #status predicted <TM2>
F/F121-142/Domain: transmembrane #status predicted <TM3>
F/F162-184/Domain: transmembrane #status predicted <TM4>
F/F208-238/Domain: transmembrane #status predicted <TM5>
F/F257-284/Domain: transmembrane #status predicted <TM6>
F/F291-314/Domain: transmembrane #status predicted <TM7>
F/F24/Binding site: carbohydrate (asn) (covalent) #status predicted
F/F119-198/Disulfide bonds: #status predicted
F/F161,253/Binding site: phosphate (Ser)(covalent) (by CAMP-dependent kinase)
F/F327/Binding site: palmitate (Cys) (covalent) #status predicted

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 67.6341 Seconds
(without alignments)
1571.954 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDGEGAREPPWP.....DTGGDTVGTETTSANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	907	42.1	363	13	O93413
2	861.5	40.0	364	11	Q8BWX8
3	851	39.5	374	13	O93412
4	825.5	38.3	366	6	Q8MH25
5	608	28.2	559	13	O93414
6	487.5	22.6	426	4	Q8NE20
7	485.5	22.5	426	4	Q9HB89
8	478.5	22.2	395	11	Q91276
9	478.5	22.2	395	11	Q8BZ39
10	477.5	22.2	395	11	Q9JTB1
11	477	22.1	415	4	Q96AM5
12	476	22.1	415	4	Q9GZ04
13	475.5	22.1	395	11	Q9ESQ4
14	457.5	21.2	405	11	O55040
15	453	21.0	477	5	Q8ITC7
16	437	20.3	418	5	O17239

17	433	20.1	428	5	Q9VFN4
18	430	20.1	430	5	Q8ITC9
19	420	19.5	595	5	Q8ITD0
20	420	19.5	599	5	Q9VFW6
21	411	19.1	402	11	Q9JTB2
22	406	18.8	412	11	Q9JTB5
23	396.5	18.4	416	11	Q8VIF5
24	395.5	18.4	416	11	Q920Q5
25	392	18.2	660	5	Q9VFW5
26	390	18.1	658	5	Q8ITD1
27	375	17.4	378	5	Q18701
28	370.5	17.2	385	11	Q9JK40
29	356.5	16.5	401	13	Q8JFZ6
30	354.5	16.5	401	13	Q9DDR0
31	350	16.2	477	13	Q8JID5
32	349.5	16.2	367	6	Q8MI04
33	341.5	15.8	370	13	Q8UWL5
34	339.5	15.8	390	13	Q8QGQ4
35	339	15.7	145	6	O97914
36	339	15.7	352	11	O88820
37	339	15.7	352	11	Q9R297
38	339	15.7	397	13	Q9DDR1
39	338.5	15.7	513	13	Q9DFA9
40	337.5	15.7	366	4	Q8IWP3
41	337.5	15.7	404	13	Q8JFZ7
42	336	15.6	352	11	Q9QMW3
43	335	15.5	382	11	Q9ERT2
44	333	15.5	397	13	Q8JFZ5
45	332	15.4	384	4	Q8IY58

ALIGNMENTS

RESULT 1

ID O93413 PRELIMINARY; PRT; 363 AA.

AC O93413; (T-EMBLrel. 08, Created)

DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)

DE Orphan G protein-coupled receptor.

OS Spherooides nephelus.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Spherooides.

OX NCBI_TaxID=39110;

RN [1]

RP MEDLINE=20092336; PubMed=10628755;

RA Palya O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,

RA Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,

RA Patchett A.A., Howard A.D., Smith R.G.;

RT "Ligand activation domain of human orphan growth hormone (GH)

RT secretagogue receptor (GHS-R) conserved from pufferfish to humans."

RL Mol. Endocrinol. 14:160-169(2000).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

DR EMBL; AF082210; AAC33473.1; -

DR InterPro; IPR000276; GPCR_Rhodopn.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.

DR PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.

KW G-protein coupled receptor; Receptor; Transmembrane.

SQ SEQUENCE 363 AA; 41324 MW; 74518BD9CD310991 CRC64;

Query Match 42.1%; Score 907; DB 13; Length 363;

Best Local Similarity 50.6%; Pred. No. 4.7e-67;

Matches 180; Conservative 56; Mismatches 76; Indels 44; Gaps 5;

OY 31 SPFPLGALVPVTAACLCLFVVGVSGNVVTVMLIGRYDRMTTNNLYLGSNAVSDLLILLG 90

Db 31 SLFPASTLIPVTICILIFVVGVTGNTWTIIIOYFKDMKTTNTNLYSSMAVSDLVFLC 90
QY 91 LPDLYLWRSRPMVFGPLCRSLVYEGECTATLHMTALSVERVLAICRPLRARVLY 150
Db 91 LPDLYLWRSRPMVFGPLCRSLVYEGECTATLHMTALSVERVLAICRPLRARVLY 150
QY 151 TRRRVRLIALVMAVALISAGPFLVLGVGEODPGISVVVGLNGTARIASSPLASSPPLWL 210
Db 151 TRRRVRLIALVMAVALISAGPFLVLGVGEODPGISVVVGLNGTARIASSPLASSPPLWL 210
QY 211 SRAPPPSPGPGPETAFAAALFSRECRPSPAQL--GALRVMLWVTYATFFLPFLCLSLY 268
Db 186 ----HPDYNTG-----QCKTGVAISSGQLHIMVSTTYFFCPMLCLFLY 229
QY 269 LIGRELWSSRPLRGPAAASGRCHROTVRVLLVVLAFIICWLPFHVGHRIIYNTEDSR 328
Db 230 SIGCKLWKSNDLQGPCALARSERSHRTVKILVVVVLAFIICWLPFHVGHRIIYNTEDSR 289
QY 329 MMVFSQYFNIVALQFLYLSASINPILYNLSKRYAAAFKLLARKSRPRGFHRSR 384
Db 290 TAMSQNFNASWVLCYLSASINPILYNLSKRYAAAFKLLARKSRPRGFHRSR 384

RESULT 2
Q8BWX8 PRELIMINARY; PRT; 364 AA.
AC Q8BWX8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Growth hormone secretagogue receptor type 1 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK049871; BAC33866.1; --
SQ SEQUENCE 364 AA; 40969 MW; 8F1214E58EF3B2E8 CRC64;

Query Match 40.0%; Score 861.5; DB 11; Length 364;
Best Local Similarity 48.3%; Pred. No. 2,7e-63;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;
QY 5 WNGSDGEGAREP-----PWPALPPCD---ERRCSPFPLGALVPVTVAVCLFLVVGVS 55
Db 2 WNAT--PSEPEPENVTLDDLWDASPGNDSLSDELPLFPAPLAGVTATCVLFFVVGIS 59
QY 56 NVVTVMILGYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPMVFGPLCRSL 115
Db 60 NLTLVVSFRFRELRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPMVFGPLCRSL 119
QY 116 YVGEGETATLLHMTALSVERVLAICRPLRARVLTTRRRVRLIALVMAVALISAGPFL 175
Db 120 FVSESCTYATVLTALSVERVLAICRPLRARVLTTRRRVRLIALVMAVALISAGPFL 179
QY 176 LVGVEODPGISVVVGLNGTARIASSPLASSPPLWLRAPPPSPGETAFAAALFSREC 235
Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQLGALRVMLWVTYATFFLPFLCLSLYLYGRELWSSRPLRGPAAASGR 290
Db 198 RATEFAVRSGLLTVMWVSSVFFLPVFCVLTVLYIGRKLW--RR--RGDAAGVSLRD 253
QY 291 RGHROTVRVLLVVLAFIICWLPFHVGHRIIYNTEDS---RMMVFSQYFNIVALQFLYLS 347

Db 254 QNHKQTVKMLVAVVPFAPILCMLPFPHVGRYLFSSKSPGSLIAQISQYCNLVSFVLFLYLS 313
QY 348 ASINPILYNLSKRYAAAFKLL 370
Db 314 AAINPILYNLSKRYAAAFKLL 336
RESULT 3
O93412 PRELIMINARY; PRT; 374 AA.
AC O93412
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=391110;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20092336; PubMed=10628755;
RA Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
RT secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF082209; AAC33472.1; --
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR000524; HTH_GntR.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 374 AA; 42324 MW; 2CF9304F004C7A16 CRC64;

Query Match 39.5%; Score 851; DB 13; Length 374;
Best Local Similarity 45.5%; Pred. No. 2.1e-62;
Matches 187; Conservative 65; Mismatches 101; Indels 58; Gaps 12;
QY 5 WNGSDGEGAREPMPALPPCDERRCSPFPLGALVPVTVAVCLFLVVGVSNNVTVMIL 63
Db 13 WEGSHNGTAGLELP-----LNYSIPLLAVITVACTVLTFTVGVGNVMTLVV 61
QY 64 GRYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPMVFGPLCRSLYVSGCTY 123
Db 62 SRYRDMRTTNLYLGSMAVSDLFIFVCMPLDLYRMWRYPWRPFGDLCALCKLQFVSSECTY 121
QY 124 ATLLHMTALSVERVLAICRPLRARVLTTRRRVRLIALVMAVALISAGPFLVGVGEODP 193
Db 122 STILCTALSVERVLAICRPLRARVLTTRRRVRLIALVMAVALISAGPFLVGVGEODP 181
QY 184 GISVVPGLNGTARIASSPL--ASSPPLWLRAPPPSPGETAFAAALFSRECRPS--PA 240
Db 182 --IMPEN-----SSDLNNESSWPL-----EAVDTRECMTOYAV 212
QY 241 QIGALRVMLWVTYATFFLPFLCLSLYLYGRELWSSRPLR--GPAASGRGRHROTVRV 299
Db 213 ESSLMEANVWLSVFFFPVFCVLTVLYIGRKLW--RR--RGDAAGVSLRD 272
QY 300 LNVVLAFLVLCWLPFHVGHRIIYNTEDS---RMMVFSQYFNIVALQFLYLSASINPILYN 356
Db 273 LNVVLAFLVLCWLPFHVGHRIIYNTEDS---RMMVFSQYFNIVALQFLYLSASINPILYN 332
QY 357 LISKYRAAAAFKLLARKSRPRGFHRSRDTAGEVAGDTGGDTVGYTETSAN 407

Db 333 TMSWKYGAVALFGVSDSPQ---RGR-TASTVKMD-----GWTESTVS 373

RESULT 4

Q8MHZ5 PRELIMINARY; PRT; 366 AA.
 AC Q8MHZ5
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Ghrelin/growth hormone secretagogue receptor.
 OS Ovis aries (Sheep).
 OC Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1, 2, and 3; TISSUE=Pituitary;
 RA Murata T., Ikegami R., Morita Y., Shinozaki K.;
 RT "Sheep Ghrelin/Growth Hormone Secretagogue Receptor";
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY093948; AAM19733.1;
 DR EMBL; AY093949; AAM19733.1;
 DR EMBL; AY093950; AAM19735.1;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR004704; IID_comp.
 DR Pfam; PF00001; 7tm.1; 1.
 DR Pfam; PF03613; EIID-AGN; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_F2_1; 1.
 KW Receptor.
 SQ SEQUENCE 366 AA; 41486 MW; 2F276651BC6C5D57 CRC64;

Query Match 38.3%; Score 825.5; DB 6; Length 366;
 Best Local Similarity 42.3%; Pred. No. 2.6e-60;
 Matches 176; Conservative 62; Mismatches 113; Indels 65; Gaps 8;
 QY 5 WNGSDGPE-----GAREPPHPALPPCD---ERRCSPEPLGALVPTAVCLCLFVVGSGN 56
 DB 2 WNAITRSEELGNLTLPDLDDAAPDNDLSLTDELPLFPAPLLAGVATATCVALLFVVGAGN 61
 QY 57 VVTVMILGRYDMRTTNLVLGSMVSDLLILGLPDLVLRMRSPWVFGPLLCRLSL 116
 DB 62 LLLTMLVVSRELTNTNLLSSNAFSDLLIFLCMLDVLVLRHYPWNLGDLCKLFQF 121
 QY 117 VSGCTVATLLHMTALSVRYLAICRPLRVLVTRRRRRLAIVLWAVALLSAGPFLFL 176
 DB 122 VSESCTVASVLITALSVERYFAICFPLRAKVITKGRVLAIVLWAVAFCSAWPIFML 181
 QY 177 VGVQDQPGISVVGPLNGTARIASSPLASSPPLMSRAPPPSPGPTAAALAFSRECR 236
 DB 182 VGVHE-----NGT-----DPRD-----TNECR 199
 QY 237 PS--PAQLGALRVMLVTTAYFFLPFLCLSLYLGLRELWSSRRPLRGAASGRGRHR 294
 DB 200 ATEFAVSGLLTINWVSSIFFFLPVCLTVLYSLIGRKLWRRSRSEVVGASURDNQHK 259
 QY 295 QTVRVLVVLVLAFTICLPHVGRYIYNTEDS---RMVFSQYFNIVALQLFYLASIN 351
 DB 260 QTVKMLAVVFAFVLCWLPFHVGRYLSKSPGVSVEIAQISQYCNLVSVFLFFYFAAIN 319
 QY 352 PILYNLSKKYRAAFAKLLARKSRPRGRFHSRTAGEVAGDGTGDTGVGTETSAN 407
 DB 320 PILYNMSKKYRAVAVFKLLGPFPPFSQKSLTKDESSR-----AWTESSIN 365

RESULT 5

Q93414 PRELIMINARY; PRT; 559 AA.
 ID Q93414
 AC Q93414;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Orphan G protein-coupled receptor.
 OS Spherooides nephelus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Spherooides.
 OC NCBI_TaxID=39110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20092336; PubMed=10628755;
 RA Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
 RA Gao Y.D., Schleich K.D., Yang L., Morriello G.J., Nargund R.,
 RA Patchett A.A., Howard A.D., Smith R.G.;
 RT "Ligand activation domain of human orphan growth hormone (GH)
 RT secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
 RL Mol. Endocrinol. 14:160-169(2000).
 DR EMBL; AF082211; AAC33474.1; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_F2_1; 1.
 KW Receptor.
 SQ SEQUENCE 559 AA; 61956 MW; AFB9F35FAP0CDSF6 CRC64;
 Query Match 28.2%; Score 608; DB 13; Length 559;
 Best Local Similarity 29.7%; Pred. No. 4.3e-42;
 Matches 156; Conservative 65; Mismatches 124; Indels 180; Gaps 12;
 QY 22 LPPCDERRC---SPFPLGALVPTAVCLCLFVVGSGNVVTVMLIGRYRDMRTTNLYLG 78
 DB 20 LHKCSNQECHWEEPV-FGMIVCVTIYIPLMLGLLGNILTLVWLRPWRSTYLYLS 78
 QY 79 SMAVSDLLILLGLPDLVLRMRSPWVFGPLLCRLSLYLGVGCTYATLLHMTALSVRYL 138
 DB 79 SLAVSDILILLGLPDLVLRMRSPWVFGPLLCRLSLYLGVGCTYATLLHMTALSVRYL 138
 QY 139 AICRPLRVLVTRRRRRLAIVLWAVALLSAGPFLVLCVE-----QDPGIS 186
 DB 139 AVCWPTAKTVVTRRTTIIGICILWGAALSAAPVVMVGVSEVQSDQGLSGWRSGAW 198
 QY 187 -----VVPGL----- 191
 DB 199 TGKEGKGFIFGRERENDKGLKDSLEBMMWKMMECCGDKNGVTGFKGDKSLEV 258
 QY 192 -NGTARIASSPLASSPPLMSRAPPPSPGPE-----TAAALAFSRECRPS--PAQLGA 244
 DB 259 GECTKEQHEGARAAGEEAQNMKEDEGGGGGGGDDTGGGRMQVDTRECRCTDYAVSSGL 318
 QY 245 LRVLWVTTAYFFLPFLCLSLYLGLRELWSSRRPLRGAASGRGRGRHQTQVRLVVLV 304
 DB 319 LSAMLVLSNMVFLVPLVCLGLVGLYGLRTLW-----LRS-QISRRDNNNTVRLGLVIV 372
 QY 305 LAFLICWLPFHVGRYIY-----INTE-----DSR----- 328
 DB 373 LVFLVCLWLPFHVGRITFFPSLGSDRPGVNAHALLSRVLPLELPPPGALGESDEAAGDAP 432
 QY 329 ----- 328
 DB 433 SEATGRGDGRAVLDTRGTARSDEGAAGSPSTPEATADYDAENSTPLDDTHSHSQYF 492
 QY 329 MMVFSQYFNIVALQLFYLASINPILYNLSKKYRAAFAKLLAR 373
 DB 493 LYLVSQYFNVLVSVLFLYLAALINPLYNLSMSRYRHAVHSLPRR 537
 RESULT 6
 Q8NE20
 ID Q8NE20
 AC Q8NE20;
 PRELIMINARY; PRT; 426 AA.

DT	01-OCT-2002 (TReMBUrel. 22, Created)	
DT	01-OCT-2002 (TReMBUrel. 22, Last sequence update)	
DT	01-MAR-2003 (TReMBUrel. 23, Last annotation update)	
DE	G protein-coupled receptor 66.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
NCBI_TaxID=9606;		
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE=Testis;	
RC	Straubeberg R.;	
RA	Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; BC036543; AAH36543.1; -.	
DR	Genew; HGNC:4518; GPR66.	
DR	InterPro; IPR000276; GPCR_Rhodpsn.	
DR	InterPro; IPR005390; NeuroMedinU.	
DR	InterPro; IPR005391; NeuroMedinU.	
DR	Fram; PF00001; 7tm_1; 1.	
DR	PRINTS; PR00237; GPCRHHODOPSN.	
DR	PRINTS; PR01565; NEUROMEDINUR.	
DR	PRINTS; PR01566; NEUROMEDINUR.	
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.	
DR	PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.	
KW	Receptor.	
QY	SEQUENCE 426 AA; 47322 MW; 7BFB5DE2E2DC686 CRC64;	
QY	Query Match 22.6%; Score 487.5; DB 4; Length 426;	
QY	Best Local Similarity 33.7%; Pred. No. 3e-32;	
QY	Matches 128; Conservative 64; Mismatches 123; Indels 67; Gaps 10;	
QY	39 PVPTAVCLCLFVVGSGVGNVTVMLIGRYDRMTTNNLYLGSMAVSDLLILL-GLPFOLYR 97	
Db	61 MPICATYLLIFVVGAVGNGLTCLVLRHKAMRTPTNYLPSLAVSDLLVLLVGLPLEIYE 120	
QY	98 LWRSRPWFVGPGLLCRLSLYVGGCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRA 157	
Db	121 MWHNYPFLGVCYCFRTLLFEMVCLASVLNVTALSVERYVAVVHPLOARSMVTRAHVRR 180	
QY	158 LIAVLMAVALISAGPFLFLVGVQEDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPS 217	
Db	181 VLGVAGLMLCSPNTSLHQIQO---LHV----- 207	
QY	218 PPSGPETAEEAALPSRECRSPSAQGLARVLMWTTA--YFPLPFLCISILYGLIGRELW 275	
Db	208 PCRGVPDPSAVCMVLRP-----RALYNNVQTALLPFLPMAIMSVLYLLIGLELR 259	
QY	276 SSRRLP-----RQPAAS-----GRGRHGQTVRVLVVVLAFIICMLPHVHGRII 320	
Db	260 RERLLMQEAKGRGSAARSRYTCLQOHDRGRQVTQMLFVLVVVFGICWAPFHADRV 319	
QY	321 Y--INTEDSRMYPYSQFNVIALQLFYLSASINPILNLSIKKYRAAAFKLL----LARK 374	
Db	320 WSVVSWQTDGLHLAFQHHVVISGIFFGYLSAANPVLVSLMSSRRPRTFQEALCLGACCHR 379	
QY	375 SRPR--GFHRSRDTAGEVAGDTG 395	
Db	380 LRPRHSSHLSRWTTGSLCDVG 402	
RESULT 7		
Q9HB89	PRELIMINARY; PRT; 426 AA.	
ID	Q9HB89	
AC	Q9HB89; O43664;	
DT	01-MAR-2001 (TReMBUrel. 16, Created)	
DT	01-MAR-2001 (TReMBUrel. 16, Last sequence update)	
DE	01-MAR-2003 (TReMBUrel. 23, Last annotation update)	
DE	NeuroMedin U receptor 1 (Orphan G protein-coupled receptor).	
GN	NMUR1	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
NCBI_TaxID=9606;		

```
[1] SEQUENCE FROM N.A.  
RP MEDLINE=2049068; PubMed=10899166;  
RX Radatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,  
RA Botetu L.W., Zhou S., Kouranova E.V., Nagorny R., Guevara M.S.,  
DA Dai M., Lerman G.S., Vayssie P.J., Branchek T.A., Gerald C., Forrest C.,  
RA Adam N.;  
RT "Identification and Characterization of Two Neuromedin U Receptors  
Differentially Expressed in Peripheral Tissues and the Central Nervous  
System.";  
RL J. Biol. Chem. 275:32452-32459(2000).  
  
[2]  
RN SEQUENCE OF 24-426 FROM N.A.  
RP MEDLINE=99000845; PubMed=9782091;  
RX Tan C.P., McKee K.K., Liu Q., Palyha O.C., Feighner S.D.,  
RA Hreniuk D.L., Smith R.G., Howard A.D.;  
RT "Cloning and characterization of a human and murine T-cell orphan G-  
protein-coupled receptor similar to the growth hormone secretagogue  
and neurotensin receptors.";  
RL Genomics 52:223-229(1998).  
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
DR EMBL; AF272362; AAC24793.1; .  
DR EMBL; AF044601; AAC02680.1; .  
DR EMBL; AF044600; AAC02680.1; JOINED.  
DR HSPB; P02699; IPF88.  
DR InterPro; IPR000276; GPCR_Rhodpsn.  
DR InterPro; IPR005390; NeuroMedinu.  
DR InterPro; IPR005391; NeuromedinU.  
DR Pfam; PF00001; 7tm_1; 1.  
DR PRINTS; PR00237; GCRRHODOPSN.  
DR PRINTS; PR01565; NEUROMEDINUR.  
DR PRINTS; PR01566; NEUROMEDINUR.  
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.  
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_2; 1.  
SQ G-protein coupled receptor; Receptor; Transmembrane.  
KW SEQUENCE 426 AA; 47350 MW; F8DD781C020F04AA CRC64;
```

Query Match 22.5%; Score 485.5; DB 4; Length 426;
Best Local Similarity 33.7%; Pred.No 4.4e-32;
Matches 129; Conservative 63; Mismatches 124; Indels 67; Gaps 10;

Qy		39	PVTVAVCLCFVVGVSGNVVTMLIGRYRDMRTTTNLYLGSMASVDLLILL-GLPFDLYR	97
Dd	:	:	: :: :	:
Yq	:	:	:: :	:
Dd	:	61	MPICATVLLIFVVGAVGNGITCLVLIRHKAMRPTNYLFLSAVSDLLLVLLGLPLEYE	120
Yq	:	:	:: :	:
Dd	:	98	LWSRPPWFGLLCRLSLYGVECTATLHLMTALSVERYLAI CRPLRARVLVTRRVRA	157
Dd	:	:	:: :	:
Dd	:	121	MWNIYPFLGGVCYCFTRTLFEWCVLASVINLTALSVERYVAHVHPLOARSVMVTRAHR	180
Yq	:	:	:: :	:
Yq	:	158	LIAVLMAVALLSAGPFLLFGVGDPGISVVPGLNGTARTIASPLASSPPLWLSRAPPS	217
Dd	:	:	:: :	:
Dd	:	181	VLGAVGLAWCLSIPNTSLHGIIQQ----	207
Yq	:	:	:: :	:
Yq	:	218	PPSGPETAEAAAALFSRECRPSPAQLGALRMVLWTITA--YFFLFELCSILYGLIGRELW	275
Dd	:	:	:: :	:
Dd	:	208	PCRGPVDSACVMLVRP-----RALYNMVQTALLFCFLPMAIMSVLYLLIGLR	259
Yq	:	:	:: :	:
Yq	:	276	SSRRPL-----RGPAAS-----GRERGHROTIVULLLVVLAIFI CWLFPHVGRRII	320
Dd	:	:	:: :	:
Dd	:	260	RERLLLMQEAKRGSAARSRYTCRLQHQRGRQVTKMLFVLVVFVFGICWAFPHADRVM	319
Yq	:	:	:: :	:
Yq	:	321	Y----INTSDSRMYFSCYFNIVALQLFVLSAINPIYLNLISKKYRAAFAKL---	374
Dd	:	:	:: :	:
Dd	:	320	WSVVSWNTDGLHIALFOHHVVISGIFFYLGSAANPNVILYSLMSSRFRETQEALCLGACC	379
Yq	:	:	:: :	:
Yq	:	375	SRRP---GFHSRDRTAGEVADTG	395
Dd	:	:	:: :	:
Dd	:	380	LRPHSHSLSRMWTGTSTLCDVG	402

RESULT 8
Q91276

RESULT 8
Q91Z76

```

ID Q91Z76 PRELIMINARY; PRT; 395 AA.
AC Q91Z76;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor type 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monema F.J. Jr.,
RA Gustafson E.L.;
RA "Characterization of murine neuromedin U R2 receptor.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AY057384; AAL26695.1; -.
DR InterPro; IPR000276; GPCR Rhodopsn.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005392; NeuromedinU2.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCR RHODOPSIN.
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1567; NEUROMEDINUR.
DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
DR PROSITE; PS0262; G PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44844 MW; 0BB540024566903D CRC64;

Query Match 22.2%; Score 478.5; DB 11; Length 395;
Best Local Similarity 31.5%; Pred. No. 1.5e-31;
Matches 115; Conservative 79; Mismatches 112; Indels 59; Gaps 10;

QY 30 CSPFPLGALVPVTAVCCLFVVGSGNVVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
DB CGPKRSDLSLPVSVVYALIFVVGIGNLLVCLVIARHQTLLKPTNYLFLSLAVSDLLVLL 88
QY 89 LGLPFDLYRLWRSPVPGPLLCRLSLYVGEGCTYATLLHMTALSVERYLAICPLRARV 148
DB LGMPLEYELHNYVFLFGVGCYFKTALFETVCFASILSVTSIERYVAIVHPFRAKL 148
QY 149 LVTRRRVRLAIVLWAVALLSAGPFLVGV--EQDPCISVWVGLNGTARIASSPLASSP 206
DB ESTERRALRLSLVWSVSVFSLPNTSINGIKFQFPNGSVFG-----SATCTVTK 200
QY 207 PLMLSRAPPPSPGPGTAEAAALFSRECRPSAQLGALRVMLWVTYATFF-LPFLCLSI 265
DB 201 PMV-----YNIQATSFLLPMTLSV 226
QY 266 LYGLIGRELSSRRPLRGPAASGRGRHRTV-RVLLVVLAFIICWLPFHVGRIIYINT 324
DB LYLMGLURLKRDESLEADKVTNTHRPSKSVTKMLFVLVFAICWTPPHVDRLPFSF 285
QY 325 EDSRMVFSQYFN---IVALQLFYLASINPILNLSKRYRAAFKLLA-----RK 374
DB 286 VDEWTESLAAVNLHVGVGFYLSAVNPILNLSRRFR-AAFRNVVSPCKWCHPQ 344
QY 375 SRPRG 379
DB 345 HRPOG 349

RESULT 9
Q8BZ39 PRELIMINARY; PRT; 395 AA.
AC Q8BZ39;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor type 2.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monema F.J. Jr.,
RA Gustafson E.L.;
RA "Characterization of murine neuromedin U R2 receptor.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AY057384; AAL26695.1; -.
DR InterPro; IPR000276; GPCR Rhodopsn.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005392; NeuromedinU2.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCR RHODOPSIN.
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1567; NEUROMEDINUR.
DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
DR PROSITE; PS0262; G PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44844 MW; 0BB540024566903D CRC64;

Query Match 22.2%; Score 478.5; DB 11; Length 395;
Best Local Similarity 31.5%; Pred. No. 1.5e-31;
Matches 115; Conservative 79; Mismatches 112; Indels 59; Gaps 10;

QY 30 CSPFPLGALVPVTAVCCLFVVGSGNVVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
DB CGPKRSDLSLPVSVVYALIFVVGIGNLLVCLVIARHQTLLKPTNYLFLSLAVSDLLVLL 88
QY 89 LGLPFDLYRLWRSPVPGPLLCRLSLYVGEGCTYATLLHMTALSVERYLAICPLRARV 148
DB LGMPLEYELHNYVFLFGVGCYFKTALFETVCFASILSVTSIERYVAIVHPFRAKL 148
QY 149 LVTRRRVRLAIVLWAVALLSAGPFLVGV--EQDPCISVWVGLNGTARIASSPLASSP 206
DB ESTERRALRLSLVWSVSVFSLPNTSINGIKFQFPNGSVFG-----SATCTVTK 200
QY 207 PLMLSRAPPPSPGPGTAEAAALFSRECRPSAQLGALRVMLWVTYATFF-LPFLCLSI 265
DB 201 PMV-----YNIQATSFLLPMTLSV 226
QY 266 LYGLIGRELSSRRPLRGPAASGRGRHRTV-RVLLVVLAFIICWLPFHVGRIIYINT 324
DB LYLMGLURLKRDESLEADKVTNTHRPSKSVTKMLFVLVFAICWTPPHVDRLPFSF 285
QY 325 EDSRMVFSQYFN---IVALQLFYLASINPILNLSKRYRAAFKLLA-----RK 374
DB 286 VDEWTESLAAVNLHVGVGFYLSAVNPILNLSRRFR-AAFRNVVSPCKWCHPQ 344
QY 375 SRPRG 379
DB 345 HRPOG 349

RESULT 10
Q9JIB1 PRELIMINARY; PRT; 395 AA.
AC Q9JIB1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor 2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.;
RT "Identification of receptors for neuromedin U and its role in feeding.";
```

RL Nature 406:70-74(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AF242875; AAF82756.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR005390; NeuromedinU.
DR InterPro: IPR005392; NeuromedinU2.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PRO0237; GPCRHHODOPS.
DR PRINTS: PRO1565; NEUROMEDINUR.
DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1_2; 1.
DR PROSITE: PS0262; G-PROTEIN RECEPTOR; Transmembrane.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;

Query Match 22.2%; Score 477.5; DB 11; Length 395;
Best local similarity 30.8%; Pred. No. 1.9e-31;
Matches 112; Conservative 81; Mismatches 114; Indels 57; Gaps 9;

QY 30 CSPFPGALVPTAVCLCLFVGVSGNVVTVMLIGRYDRMTTTLNLYLGSMAVSDLLI-L 88
DB 29 CGPKRSLSPVSVAVALIFLVGMGNLLVCMVLRHQTUKTPNYLFLSLAVSDLLVLL 88
QY 89 LGLPFDLYRLWRSRPWVFGPGLLCRLSLYVGEGETYATLLHMTALSVERYLAICRPLRARV 148
DB 89 LGMPLEIYEMWHNYPFLFGPGVCYKFTALFETVCFASILSVTVSVERYVAIVHPFRAKL 148
QY 149 LVTRRRVRLIALVWALLSAGPFLFLVGV--EODPGISVVPGLNLTARIASPLASSP 206
DB 149 ESTRRRALRILSLWSFSVFLPNTSIHGKIFHPFGSLVPG-----SATCTVTIK 200
QY 207 PLMLSRAPPPSPGPETAEEAALFSRECRPSPAQLGALRVMLVTVTTAYFF-LPFLCLSI 265
DB 201 PMWV-----YNIITQATSFLLPMTLISV 226
QY 266 LYGLIGRELWSSRRPLRGPAASGRGRHRTQV-RVLLVVLVLAFLICWLPFHVGRIIYINT 324
DB 227 LYLLMGLRLKRDSELEANKAVANIHRSKSVTKMLFVLVLFVFAICWTPFHVDRLFFSFV 286
QY 325 ED--SRMYFSQVFNIVALQLFYLSASINPILYNLSKKYRAAFKULLA-----RKS 375
DB 287 EEWTESLAAVFNLIHVGVGVFFYLSSAVNPILYNLSRRFR-AAFRNVVSPCKWCHPRH 345
QY 376 RPRG 379
DB 346 RPQG 349

RESULT 11
Q96AM5
ID Q96AM5 PRELIMINARY; PRT; 415 AA.
AC Q96AM5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor 2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: BC016938; AAHL6938.1; -
DR Genew; HGNC:16454; NMU2R.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR005390; NeuromedinU.
DR InterPro: IPR005392; NeuromedinU2.

DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PRO0237; GPCRHHODOPS.
DR PRINTS: PRO1565; NEUROMEDINUR.
DR PRINTS: PRO1567; NEUROMEDINUR2.
DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 415 AA; 47770 MW; 30BFEDD706436AB9 CRC64;

Query Match 22.1%; Score 477; DB 4; Length 415;
Best local similarity 31.8%; Pred. No. 2.2e-31;
Matches 115; Conservative 81; Mismatches 106; Indels 60; Gaps 11;

QY 30 CSPFPGALVPTAVCLCLFVGVSGNVVTVMLIGRYDRMTTTLNLYLGSMAVSDLLI-L 88
DB 37 CGPRRSHFFLPVSVVVPFVGVGIVNLVCLVILQHQAMKTPNYLFLSLAVSDLLVLL 96
QY 89 LGLPFDLYRLWRSRPWVFGPGLLCRLSLYVGEGETYATLLHMTALSVERYLAICRPLRARV 148
DB 97 LGMPLEIYEMWHNYPFLFGPGVCYKFTALFETVCFASILSVTVSVERYVAIVHPFRAKL 156
QY 149 LVTRRRVRLIALVWALLSAGPFLFLVGV--EODPGISVVPGLNLTARIASPLASSP 206
DB 157 QSTRRALRILGTVGFSVFLPNTSIHGKIFHPFGSLVPG-----SATFTVIK 208
QY 207 PLMLSRAPPPSPGPETAEEAALFSRECRPSPAQLGALRVMLVTVTTAYFF-LPFLCLSI 265
DB 209 PMWV-----YNIITQATSFLLPMTVISV 234
QY 266 LYGLIGRELWSSRRPLRGPAASGRGRHRTQV-RVLLVVLVLAFLICWLPFHVGRIIYINT 324
DB 235 LYLLMALRLKDKRLEADENAGNIQRCRKSVMKMLFVLVLFVFAICWAPPHIDRLFFSFV 294
QY 325 ED--SRMYFSQVFNIVALQLFYLSASINPILYNLSKKYRAAFKULLARSKPRGF 380
DB 295 EEWSESL--AAVENLVHVGVGVFFYLSSAVNPILYNLSRRFQ-AAFQNVIS-----SF 345
QY 381 HR 382
DB 346 HK 347

RESULT 12
Q9GZQ4
ID Q9GZQ4 PRELIMINARY; PRT; 415 AA.
AC Q9GZQ4; Q9NPA6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor 2 (Neuromedin U receptor-type 2) (G protein-coupled receptor TGR-1).
DE coupled receptor TGR-1).
GN NMUR2 OR NMUR2 OR TGR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.B., Artymyshyn R., Bonini J.A., Borowsky B.,
RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA Dai M., Lerman G.S., Vaysse P.J., Branchek T.A., Gerald C., Forray C.,
RA Adham N.;
RT Identification and Characterization of Two Neuromedin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System.";
RL J. Biol. Chem. 275:32452-32459(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11010960;
RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
RA Gustafson E.L., Monsma F.J. Jr., Hedrick J.A.;
RT Identification of a Novel Neuromedin U Receptor Subtype Expressed in

the Central Nervous System."; J. Biol. Chem. 275:39482-39486 (2000).
 [3] SEQUENCE FROM N.A.
 RA Pang L., Wang S., Laz T., Hedrick J.A.;
 RP Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 [4] SEQUENCE OF 4-415 FROM N.A.
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
 RA Stair J.N., Yu H., Jiang Q., Clemente M.K., Tan C.P., McKee K.K.,
 RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austen C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 RT feeding."; Nature 406:70-74 (2000).
 [5] SEQUENCE OF 4-415 FROM N.A.
 RX PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 RT Neuromedin U Receptor."; J. Biol. Chem. 275:29528-29532 (2000).
 RL J. Biol. Chem. 275:29528-29532 (2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AF272363; AAG24794.1; -;
 DR EMBL: AF292402; AAG03064.1; -;
 DR EMBL: AF242874; AAR82755.1; -;
 DR EMBL: AB041228; BAB13721.1; -;
 DR InterPro: IPR005390; NeuromedinU.
 DR InterPro: IPR005392; NeuromedinU2.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS: PRO0237; GPCRHHODOPS.
 DR PRINTS: PRO1565; NEUROMEDINUR.
 DR PRINTS: PRO1567; NEUROMEDINUR.
 DR PROSITE: PS0237; G-PROTEIN RECP F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 415 AA; 47725 MW; C2BACD84B313390F CRC64;
 Query Match 22.1%; Score 476; DB 4; Length 415;
 Best Local Similarity 31.8%; Pred. No. 2.6e-31;
 Matches 115; Conservative 81; Mismatches 106; Indels 60; Gaps 11;
 QY 30 CSPFPLGALVPVTAVALCLFVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
 DB 37 CGPRSHFFLPVSVVYVPIFVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 96
 QY 89 LGLPFDLYRLWRSRPWVFGPGLLCRLSLYVGECCGYATLLHMTALSVERYLAICRPLR 148
 DB 97 LGMPLEIYEMWNNYVPLFGVGVGYKFTALFETVCFSILSTTVSVRYVAIHPFRKL 156
 QY 149 LVTRRRRLIAVLWAVALLSAGPFLVGV--QDPGISVVPGLNGTARIASPLASSP 206
 DB 157 QSTERRALRLIGVWGSVLSFSLPNTSIHGKIFHYFNGSLVFG-----SATCTVIK 208
 QY 207 PLWSRAPPPSPSGPETAEAAALFSRECRPSAQLGALRVMLWVTYATFF-LPFLCLSI 265
 DB 209 PMWI-----YNFIQVTSFLFLLPMTVISV 234
 QY 266 LYGLIGRELWSSRRPLRGPAASGRGRHRTV-RVLLVVVLAFLIICWLPFHVGRIIYINT 324
 DB 235 LYLMALRLKKDSLEADGNANTQRCRKSVMKMLFVLVLAICWAPPHIDRLFFSFV 294
 QY 325 ED--SRMYFSQYFNIVAL---QLFYLASINPIYLNLSIKKYRAAFAKLLARKSRPRGF 380
 DB 295 EEWSESL--RAVFNLVHVSVGVFFYLSAVNPPIYLNLSRRFQ-AAFQNVIS-----SP 345
 QY 381 HR 382

Db 346 HK 347
 RESULT 13
 Q9ESQ4 PRELIMINARY; PRT; 395 AA.
 ID Q9ESQ4
 AC Q9ESQ4;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE G-protein-coupled receptor TGR-1.
 GN TGR-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20449029; PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 RT Neuromedin U Receptor."; J. Biol. Chem. 275:29528-29532 (2000).
 RL J. Biol. Chem. 275:29528-29532 (2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AB041229; BAB13722.1; -;
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR005390; NeuromedinU.
 DR InterPro: IPR005392; NeuromedinU2.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS: PRO0237; GPCRHHODOPS.
 DR PRINTS: PRO1565; NEUROMEDINUR.
 DR PRINTS: PRO1567; NEUROMEDINUR.
 DR PROSITE: PS0237; G-PROTEIN RECP F1_1; 1.
 DR PROSITE: PS0262; G-PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 395 AA; 44756 MW; EABD6D36E9C355DA CRC64;
 Query Match 22.1%; Score 475.5; DB 11; Length 395;
 Best Local Similarity 31.5%; Pred. No. 2.7e-31;
 Matches 108; Conservative 75; Mismatches 111; Indels 49; Gaps 7;
 QY 30 CSPFPLGALVPVTAVALCLFVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
 DB 29 CGPRSDLSLPVSAVALFLVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
 QY 89 LGLPFDLYRLWRSRPWVFGPGLLCRLSLYVGECCGYATLLHMTALSVERYLAICRPLR 148
 DB 89 LGMPLEIYEMWNNYVPLFGVGVGYKFTALFETVCFSILSTTVSVRYVAIHPFRKL 148
 QY 149 LVTRRRRLIAVLWAVALLSAGPFLVGV--EQDPGISVVPGLNGTARIASPLASSP 206
 DB 149 ESTRRRALRLISLVWGSVLSFSLPNTSIHGKIFHYFNGSLVFG-----SATCTVTK 200
 QY 207 PLWSRAPPPSPSGPETAEAAALFSRECRPSAQLGALRVMLWVTYATFF-LPFLCLSI 265
 DB 201 PMWV-----YNLIQVTSFLFLLPMTVISV 226
 QY 266 LYGLIGRELWSSRRPLRGPAASGRGRHRTV-RVLLVVVLAFLIICWLPFHVGRIIYINT 324
 DB 227 LYLMGLRLKDSLEANKVAVNHNHPSKSVTKMLFVLVLAICWTPPHVDRLFFSFV 286
 QY 325 ED--SRMYFSQYFNIVALQOLFYLASINPIYLNLSIKKYRAA 365
 DB 287 EEWTESLAAVFNLIHVSVGVFFYLSAVNPPIYLNLSRRPRAA 329
 RESULT 14
 O55040
 ID O55040 PRELIMINARY; PRT; 405 AA.

```
AC OS5040;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Orphan G protein-coupled receptor.
GN GPR66.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA McKee K.K., Tan C.T., Liu J., Palyha O.C., Feighner S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF044602; AAC02681.1; -.
DR MGD; MGI:1341898; Gpr66.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005391; NeuromedinU.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCRHHODOPSN.
DR PRINTS; PRO1566; NEUROMEDINUR.
DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ
Query Match 21.2%; Score 457.5; DB 11; Length 405;
Best Local Similarity 32.2%; Pred. No. 8.6e-30;
Matches 133; Conservative 62; Mismatches 141; Indels 77; Gaps 14;
QY 19 WPALPP-----CDE-----RRCSPPFLGALVPTAVCLCLFVVGSGNVVTVMLIGRVDMR 70
DB 10 WPYQPEDLNLTDALRLKYLGPQOMKQFVPCVLYLLIFVVGTLGNGLTCTVILRNKTM 69
QY 71 TTNLYLGSMAVSDLLILL-GLPDLVLRWSRPWVFGPLLCLRLSLVVGSGCTYATLLHM 129
DB 70 TPTNFFLFSVSDMLVLLVGLPLELLEMOQNPFFQLGSACVFRILLLETVCLSVLNV 129
QY 130 TALSVERYLAICRPLRARVLTTRRRVRLALIAVMAVALLSAGPFLFVVGVEQDPGISVWP 189
DB 130 TALSVERYVAVRPLQAKSVNTRAHVRYMGAIWVLTATLSLNTSLHLSQ-----LTVP 185
QY 190 GLNGTARIASSPLASSPPLWLSRAPPPSPGPETAFAALFGRCPSPQAQLGALRVML 249
DB 186 -----CRGFVFD-----SAICS-----LVGPMDFVKL-VVL 210
QY 250 WVTAYFPFLCLSLIYGLIGRELWSSRRPL-----RGPAASG-----RRGHR 294
DB 211 TTALLFFCLPMVTVISVLYLIGLRRLRRMLLOVEVGRKTAQTQETSHRRIQLQDRGR 270
QY 295 QTVRVLVAVVLAFLICMLPFHVGRIIYI-----NTDSRMMYFSQYFNIVALQLFVLSAS 349
DB 271 QVTGMFLAVVVGICWAPHADRMISLVYGHSTEGHLAY--QCWHIASGIFVYLGSA 328
QY 350 INPILYNLISKYRAAFKLL-----LARKSPRGPHRS-RDTAGVAGDTG 395
DB 329 ANPVLYSLMSTRFETFLQALGLGTQCCHRRQPVGHSHNIRLTGTSLCDVG 381
```

RESULT 15

Q8ITC7

ID Q8ITC7

AC Q8ITC7

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative Cap2b receptor.

GN CG14575.

```
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Cantons;
RA Park Y., Kim Y.J., Adams M.E.;
RA MEDLINE=22177201; PubMed=12177421;
RT "Identification of G protein-coupled receptors for Drosophila PRXamide
RT peptides, CCAP, corazonin, and AKH supports a theory of ligand-
RT receptor coevolution."
RL Proc. Natl. Acad. Sci. U.S.A. 99:11423-11428 (2002).
DR EMBL; AF522193; AAN10046.1; -.
KW Receptor.
SQ
Query Match 21.0%; Score 453; DB 5; Length 477;
Best Local Similarity 29.7%; Pred. No. 2.4e-29;
Matches 117; Conservative 68; Mismatches 135; Indels 74; Gaps 11;
QY 8 SDGEGAREPPWALPPCDERRCSP-----FPLGALVPTAVCLCLFVVGSG 55
DB 30 SDPSHGFGEEDY----ACGTFCNSKPEFVAVLGPQTLPLYKAVLITIIFFGIFITGVG 85
QY 56 NVTVMLIGRYDRMRTTNLYLGSMAVSDLLILL-GLPDLVLRWSRPWVFGPLLCLRLS 114
DB 86 NLLVCVILRHSAHMTATNYLFSVSLVSDLLYLLFGLPTEVFLYMHQYDPLFGMPFKIR 145
QY 115 LVVGEGCTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLALIAVMAVALLSAGPFL 174
DB 146 AFISEACTVSVFTIVAFSMERFLAICHPLHLVAMVGFKAIRIITALTIVTSISAIPF- 204
QY 175 FLVGVEQDPGISVVGNGTARIASSPLASSPPLWLSRAPPPSPGPETAFAALFSRE 234
DB 205 -----GLSDIQYINYPDHS-----RIEESAF----- 227
QY 235 CRPSPAQLGALRVMLVTTAYRPLPFLCLSLIYGLIGRELWSSRRPLRG--PAASGRERG 292
DB 228 CSMSPKIVNEIPVFEVSFCIFFVPMILILLYGRMGAKIRSTNKLGVQOQTNNRETR 287
QY 293 HRQ-----TVRVLVAVVLAFLICMLPFHVGRIIYIINTDSRMMYFSQYFNI-----VAL 341
DB 288 NSQMRKKTIVIRMLAAVITFFVCWCFPHLQRLIFLYAKN-----MDNYLDINEALFSIAG 342
QY 342 QLPYLSASINPILYNLISKYRAAFKLLARKS 375
DB 343 FAYTVSCTVNPVIVYSMSRRYR-VAFRELICGKA 375
```

Search completed: January 1, 2004, 06:22:26
Job time : 71.6341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 18:55:58 ; Search time 5295.11 Seconds
(without alignments)
10739.032 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atggggcagcccttggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rnd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1161	83.5	1161	6	AX154591	AX154591 Sequence
2	1078	77.6	1239	6	AX154589	AX154589 Sequence
3	1078	77.6	1239	6	AX549187	AX549187 Sequence
4	1078	77.6	1239	6	AX572965	AX572965 Sequence
5	1052	75.7	2040	9	AF034632	AF034632 Homo sapi
6	1052	75.7	163284	9	AL137000	AL137000 Homo sapi
7	1052	75.7	341560	2	AL596304	AL596304 Homo sapi
8	1052	75.7	349980	6	AX711879	AX711879 Sequence
9	1052	75.7	349980	6	AX739961	AX739961 Sequence
10	592.2	42.6	692	9	HS3339407	AJ339407 Homo sapi
11	572.2	41.2	692	9	HS3342408	AJ3342408 Homo sapi
12	557.4	40.1	615	9	HS3339459	AJ339459 Homo sapi
13	424.8	30.6	701	9	HS3326768	AJ326768 Homo sapi
14	307	22.1	192116	2	AC109158	AC109158 Mus muscu
15	283	20.4	283	6	AR168467	AR168467 Sequence
16	283	20.4	283	6	AR182284	AR182284 Sequence
17	283	20.4	283	6	AR303910	AR303910 Sequence
18	283	20.4	283	6	BD056681	BD056681 Galanin r
19	283	20.4	283	6	BD064652	BD064652 Human gal
20	283	20.4	283	6	BD064664	BD064664 Mouse gal
21	266.8	19.2	573	9	HS3341295	AJ341295 Homo sapi
22	242.6	17.5	1351	5	AB095997	AB095997 Gallus ga
23	242.6	17.5	1703	5	AB095996	AB095996 Gallus ga
24	242.6	17.5	1751	5	AB095995	AB095995 Gallus ga
25	242.6	17.5	4121	5	AB095994	AB095994 Gallus ga
26	234.8	16.9	1063	6	AR156351	AR156351 Sequence
27	234.8	16.9	1101	4	SSU60178	U60178 Sub scrofa
28	233.2	16.8	870	4	SSU60180	U60180 Sub scrofa
29	230	16.5	1029	6	AR156352	AR156352 Sequence
30	230	16.5	110000	2	AC095360_2	Continuation (3 of
31	230	16.5	183099	2	AC117349	AC117349 Rattus no
32	230	16.5	195131	2	AC108400	AC108400 Mus muscu
33	230	16.5	227035	2	AC116585	AC116585 Mus muscu
34	230	16.5	228796	2	AC095173	AC095173 Rattus no
35	230	16.5	250754	2	AC130771	AC130771 Rattus no
36	229	16.5	250	6	AR269796	AR269796 Sequence
37	228.8	16.5	870	6	AX548854	AX548854 Sequence
38	228.8	16.5	870	9	HSU60181	U60181 Human growt
39	228.8	16.5	1122	6	AR156354	AR156354 Sequence
40	228.8	16.5	6787	9	AP369786	AP369786 Homo sapi
41	228.8	16.5	145593	9	AC069523	AC069523 Homo sapi
42	228.8	16.5	159005	2	AC015655	AC015655 Homo sapi
43	228.8	16.4	1092	6	AR156357	AR156357 Sequence
44	228.4	16.4	1092	6	BD094703	BD094703 GH5R liga
45	228.4	16.4	1092	6	E54510	E54510 Novel polyp

ALIGNMENTS

RESULT 1	AX154591	AX154591	1161 bp	DNA	linear	PAT 23-JUN-2001
LOCUS	Sequence 11	from Patent	WO0138355.			
DEFINITION	AX154591					
ACCESSION	AX154591					
VERSION	AX154591.1	GI:14536177				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1	Sheppard, P.O., Jaspers, S.R., Deisher, T.A. and Bishop, P.D.				
AUTHORS		Method of forming a peptide-receptor complex with zaig33 and				
TITLE		therapeutic use thereof				

Best Local Similarity 89.1%; Pred. No. 4.5e-134;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

QY 1 ATGGGAGAGCCCTTGGAAACGAGCGAGCGGCCCGGAGGGGCGGGAGCGCGTGGCCCC 60
DB 1 ATGGGAGAGCCCTTGGAAACGAGCGAGCGGCCCGGAGGGGCGGGAGCGCGTGGCCCC 60
QY 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
DB 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
QY 121 GTGACCGCTGTGCGCTTGGCTTGTCTGCTGGGGGTGAGCGCAACGTTGACCGTG 180
DB 121 GTGACCGCTGTGCGCTTGGCTTGTCTGCTGGGGGTGAGCGCAACGTTGACCGTG 180
QY 181 ATGCTGATCGGGCGCTACCGGGGATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240
DB 181 ATGCTGATCGGGCGCTACCGGGGATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240
QY 241 GCCGTGTCGACCTTATCTATCTGCTCGGGCTGCGCTTGCACCTGTACCGCCTCTGGCG 300
DB 241 GCCGTGTCGACCTTATCTATCTGCTCGGGCTGCGCTTGCACCTGTACCGCCTCTGGCG 300
QY 301 TCGGGCCCTGGGTGTTGGGGCGCGCTCTCTGCGCCCTTCTGCGCGCTGCTTCTGCGG 360
DB 301 TCGGGCCCTGGGTGTTGGGGCGCGCTCTCTGCGCCCTTCTGCGCGCTGCTTCTGCGG 360
QY 361 TGCACCTTACCGCAGCTGTGTCATGACATGACCGCGCTCAGCGTTCAGCGCTACCTTGGCCATC 420
DB 361 TGCACCTTACCGCAGCTGTGTCATGACATGACCGCGCTCAGCGTTCAGCGCTACCTTGGCCATC 420
QY 421 TGGCGCCCGCTCGCGGCCCGCTTGTGTCACCGCGCGCGCGCTCGCGCGCTCATCGCT 480
DB 421 TGGCGCCCGCTCGCGGCCCGCTTGTGTCACCGCGCGCGCGCTCGCGCGCTCATCGCT 480
QY 481 GTGCTTGGGCGGTGGCGCTGTCTCTGCGCGTCCCTTCTGCTTGTGTTGTTGGGCGTCCAG 540
DB 481 GTGCTTGGGCGGTGGCGCTGTCTCTGCGCGTCCCTTCTGCTTGTGTTGTTGGGCGTCCAG 540
QY 541 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCTCTCG 600
DB 541 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCTCTCG 600
QY 601 CCTCTCGCTCTGCGCGCTCTCTGCTCTCTGCGCGCGCACCGCGCTCCCGCGCTCG 660
DB 601 CCTCTCGCTCTGCGCGCTCTCTGCTCTCTGCGCGCGCACCGCGCTCCCGCGCTCG 660
QY 661 GGGCCCGAGACCGCGAGGCGCGCGCTGTTCAGCGCGGAAATGCGCGCGAGCGCGCG 720
DB 661 GGGCCCGAGACCGCGAGGCGCGCGCTGTTCAGCGCGGAAATGCGCGCGAGCGCGCG 720
QY 721 CAGTGGGCGCGCTGTGTCATGCTGTGGGTCAACCGCTTCTTCTTCTGCGCGCTTT 780
DB 721 CAGTGGGCGCGCTGTGTCATGCTGTGGGTCAACCGCTTCTTCTTCTGCGCGCTTT 780
QY 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCGGAGTGTGGAGCAGCGCGCGCG 840
DB 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCGGAGTGTGGAGCAGCGCGCGCG 840
QY 841 CTGCGAGGCGCGCGCTCTGGGGCGGAGAGGCGCACCGCGAGACCGTCCGCGTCTG 900
DB 841 CTGCGAGGCGCGCGCTCTGGGGCGGAGAGGCGCACCGCGAGACCGTCCGCGTCTG 900
QY 901 CGTAAGTGGAGCGCGCTGTTCCAAAGAGCGCTGCTGAGTCCGCGCGCGCGGAGCC 960
DB 901 C----- 901
QY 961 GCGCAACGCTGGGTCCCTTCCCTTCTGCGCCAGCTCTGGGCGCGCTTCCAGCTCCC 1020
DB 902 ----- 901
QY 1021 TTTCCTATTTCGATTCAGGCTTCCACCGCGTGGTGGTCTTGGCAATTAATTTG 1080
DB 1021 TTTCCTATTTCGATTCAGGCTTCCACCGCGTGGTGGTCTTGGCAATTAATTTG 1080

Db 902 -----TGGTGGTGGTCTTGGCATTTATAATTTG 929
QY 1081 CTGTTGGCCCTTCCAGCTTGGCAGAAATCATTTACATAAAACACGGAAGATTCGCGGATGAT 1140
DB 930 CTGTTGGCCCTTCCAGCTTGGCAGAAATCATTTACATAAAACACGGAAGATTCGCGGATGAT 989
QY 1141 GTACTTCTCTCAGTACTTTAAACATCGTCTGCTCTGCAACTTTTCTATCTAGAGCGCATCTAT 1200
DB 990 GTACTTCTCTCAGTACTTTAAACATCGTCTGCTCTGCAACTTTTCTATCTAGAGCGCATCTAT 1049
QY 1201 CAACCAATCTCTCAACCTCATTTCAAGAGTACAGAGCGCGCGCTTTAAACTGCT 1260
DB 1050 CAACCAATCTCTCAACCTCATTTCAAGAGTACAGAGCGCGCGCTTTAAACTGCT 1109
QY 1261 GCTCGCAAGAGTCCAGGCGCGAGAGCTTCCACAGAAGCAGGAGACATCTCGGGGGAAGT 1320
DB 1110 GCTCGCAAGAGTCCAGGCGCGAGAGCTTCCACAGAAGCAGGAGACATCTCGGGGGAAGT 1169
QY 1321 TGCAGGGGACATCGAGGAGACACGCTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1380
DB 1170 TGCAGGGGACATCGAGGAGACACGCTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1229
QY 1381 GATGGGATAA 1390
DB 1230 GATGGGATAA 1239

RESULT 3
AX549187 LOCUS 1239 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 472 from Patent WO02061087.
ACCESSION AX549187
VERSION AX549187.1 GI:25813904
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burner, G.C., Roush, C.L. and Brown, J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides
JOURNAL Patent: WO 02061087-A 472 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES Location/Qualifiers
source 1..1239
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 172 a 431 c 397 g 239 t
ORIGIN

Query Match 77.6%; Score 1078; DB 6; Length 1239;
Best Local Similarity 89.1%; Pred. No. 4.5e-134;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

QY 1 ATGGGAGAGCCCTTGGAAACGAGCGAGCGGCCCGGAGGGGCGGGAGCGCGTGGCCCC 60
DB 1 ATGGGAGAGCCCTTGGAAACGAGCGAGCGGCCCGGAGGGGCGGGAGCGCGTGGCCCC 60
QY 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
DB 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
QY 121 GTGACCGCTGTGCGCTTGGCTTGTCTGCTGGGGGTGAGCGCAACGTTGACCGTG 180
DB 121 GTGACCGCTGTGCGCTTGGCTTGTCTGCTGGGGGTGAGCGCAACGTTGACCGTG 180
QY 181 ATGCTGATCGGGCGCTACCGGGGATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240
DB 181 ATGCTGATCGGGCGCTACCGGGGATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240

241 GCGGTGTCGACCTACTCATCTGCTCGGGCTGCCGTTGACCTGTACCGCCTCTGGCGC 300
 Db GCGGTGTCGACCTACTCATCTGCTCGGGCTGCCGTTGACCTGTACCGCCTCTGGCGC 300
 301 TCGGGCCCTGGGTGTTGGGGCCGCTGCTCTGCGCCCTGCTCCCTCTACGTTGGGGAGGC 360
 Db TCGGGCCCTGGGTGTTGGGGCCGCTGCTCTGCGCCCTGCTCCCTCTACGTTGGGGAGGC 360
 361 TGACCTACGCCACCTGCTGCAATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
 Db TGACCTACGCCACCTGCTGCAATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
 421 TCGCGCCGCTCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db TCGCGCCGCTCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 481 GTGCTCTGGGCGGTGGCGCTGCTCTGCGCGCTGCTCTGCGCGCTGCTCTGCTGCTGCTGCT 540
 Db GTGCTCTGGGCGGTGGCGCTGCTCTGCGCGCTGCTCTGCGCGCTGCTCTGCTGCTGCTGCT 540
 541 CAGGACCCGGCATCTCCGTAGTCCCGGCTCAATGGCACCGCGGATCGCTCTCTCG 600
 Db CAGGACCCGGCATCTCCGTAGTCCCGGCTCAATGGCACCGCGGATCGCTCTCTCG 600
 601 CCTCTCGCCTCGTCCGCGCTCTCTGCTCTGCGCGCGCACCGCGCTCCCGCGCTCG 660
 Db CCTCTCGCCTCGTCCGCGCTCTCTGCTCTGCGCGCGCACCGCGCTCCCGCGCTCG 660
 661 GGGCCCGAGACCGCGAGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db GGGCCCGAGACCGCGAGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 721 CAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db CAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 781 CTGTGCTCAGCATCTCTACGGCTCATCGCGCGGAGCTGCGGAGCAGCGCGCGCGC 840
 Db CTGTGCTCAGCATCTCTACGGCTCATCGCGCGGAGCTGCGGAGCAGCGCGCGCGC 840
 841 CTGGAGGCGCGCGCTCGGGCGGAGAGAGGCCACCGGAGACCGTCCGCGCTGCTG 900
 Db CTGGAGGCGCGCGCTCGGGCGGAGAGAGGCCACCGGAGACCGTCCGCGCTGCTG 900
 901 CGTAACTGAGAGCGCGCTGCTCCAAAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db CGTAACTGAGAGCGCGCTGCTCCAAAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 1021 TTTCTATTTGATTCAGCTCCACCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
 Db TTTCTATTTGATTCAGCTCCACCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
 1081 CTGGTTGCCCTTCCAGTTGGCAGATCAATTAACAACAGAGATTCGCGATGAT 1140
 Db CTGGTTGCCCTTCCAGTTGGCAGATCAATTAACAACAGAGATTCGCGATGAT 1140
 930 CTGGTTGCCCTTCCAGTTGGCAGATCAATTAACAACAGAGATTCGCGATGAT 989
 1141 GTACTTCTCAGTACTTAAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db GTACTTCTCAGTACTTAAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 1201 CAACCCCAATCTCTACAACTCTTTTCAAGAGAGTACAGAGCGCGCGCTTTAACTGCT 1260
 Db CAACCCCAATCTCTACAACTCTTTTCAAGAGAGTACAGAGCGCGCGCTTTAACTGCT 1260
 1050 CAACCCCAATCTCTACAACTCTTTTCAAGAGAGTACAGAGCGCGCGCTTTAACTGCT 1109
 1261 GCTCGAGAGAGTCCAGGCGGAGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1320
 Db GCTCGAGAGAGTCCAGGCGGAGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1320
 1110 GCTCGAGAGAGTCCAGGCGGAGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1169
 1321 TGACGGGACACTGGAGAGACAGCGTGGGTACACCGAGACAGCGCTAACTGAAGAC 1380

Db 1170 TGACGGGACACTGGAGGACACCGTGGGCTACCGAGACAGCGCTAACTGAAGAC 1229
 Qy 1381 GATGGGATAA 1390
 Db 1230 GATGGGATAA 1239
 RESULT 4
 AX572965
 LOCUS AX572965 1239 bp DNA linear PAT 29-NOV-2002
 DEFINITION Sequence 1 from Patent WO02057791.
 ACCESSION AX572965
 VERSION AX572965.1 GI:26005013
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Brown, J. P., Burmer, G. C., Roush, C. L. and Kulander, B. G.
 TITLE Diagnostic and therapeutic compositions and methods related to gpr
 JOURNAL Patent: WO 02057791-A 1 25-JUL-2002;
 Lifespan Biosciences, Inc. (US)
 FEATURES
 source 1..1239
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 172 a 431 c 397 g 239 t
 ORIGIN
 Query Match 77.8%; Score 1078; DB 6; Length 1239;
 Best Local Similarity 89.1%; Pred. No. 4.5e-134;
 Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
 Qy 1 ATGGGAGAGCCCTTGGAAACGGACGACGCGCCGAGGGGCGGAGCGCGCTGGCCCC 60
 Db 1 ATGGGAGAGCCCTTGGAAACGGACGACGCGCCGAGGGGCGGAGCGCGCTGGCCCC 60
 Qy 61 GCGCTGCCGCTTGGACGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGCGCTGGTCCG 120
 Db 61 GCGCTGCCGCTTGGACGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGCGCTGGTCCG 120
 Qy 121 GTACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 GTACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCTTGTACCTGGGCGAGCATG 240
 Db 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCTTGTACCTGGGCGAGCATG 240
 Qy 241 GCGGTGTCGACACTACTCATCTGCTCGGGCTGCGCTGCGACCTGACCGCTCTGGCGC 300
 Db 241 GCGGTGTCGACACTACTCATCTGCTCGGGCTGCGCTGCGACCTGACCGCTCTGGCGC 300
 Qy 301 TCGCGGCGCTGGGTGTTGCGGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 301 TCGCGGCGCTGGGTGTTGCGGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Qy 361 TGACCTACGCCACCGTGTGCAATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
 Db 361 TGACCTACGCCACCGTGTGCAATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
 Qy 421 TCGCGCGCGCTCGCGCGCGCGCTTGTGGTCAACCGCGCGCGCTCGCGCGCGCTCATCGCT 480
 Db 421 TCGCGCGCGCTCGCGCGCGCGCTTGTGGTCAACCGCGCGCGCTCGCGCGCGCTCATCGCT 480
 Qy 481 GTGCTCTGGGCGGTGGCGCTGCTCTGCGCGCTGCTCTGCGCGCTGCTCTGCTGCTGCTGCTGCT 540
 Db 481 GTGCTCTGGGCGGTGGCGCTGCTCTGCGCGCTGCTCTGCGCGCTGCTCTGCTGCTGCTGCTGCT 540

Qy 541 CAGAGCCCGGCGATCTCCGTAGTCCCGGGCTCAATGGCAACCGCGGATCGCTCTCTCG 600
Db 541 CAGAGCCCGGCGATCTCCGTAGTCCCGGGCTCAATGGCAACCGCGGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCTGTCGGCGCTCTCTGCTCTCGGGGCGCCACCGCGCTCCCGCGCTCG 660
Db 601 CCTCTCGCTCTGTCGGCGCTCTCTGCTCTCGGGGCGCCACCGCGCTCCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTTCAGCGCGCGAATGCCCGCGAGCCCGCG 720
Db 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTTCAGCGCGCGAATGCCCGCGAGCCCGCG 720
Qy 721 CAGTGGGGCGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGTGGGGCGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGGCGAGCTGTGGAGCAGCGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGGCGAGCTGTGGAGCAGCGCGCGCG 840
Qy 841 CTGCGAGCGCGCGCTCTGCGGGGCGAGAGCGCCACCGCGAGACCGCTCGCTCTG 900
Db 841 CTGCGAGCGCGCGCTCTGCGGGGCGAGAGCGCCACCGCGAGACCGCTCGCTCTG 900
Qy 901 CGTAAGTGGAGCGCGCTGTTTCAAGAAGCGCTGCTGCGAGTCCGCGCGCGCGGAC 960
Db 901 C----- 901
Qy 961 GCGCAAGCTGGGTCCCTTCCCTGCTCGCCAGCTCTGGGGCGCGCTTCAGCTCCC 1020
Db 902 ----- 901
Qy 1021 TTTCCTATTTCGATTCAGCTCCACCGCGTGGTGGTCTGCAATTTAATTTG 1080
Db 902 -----TGGTGGTGGTCTGCAATTTAATTTG 929
Qy 1081 CTGTTGCCCTTCCAGTGGCAGATCAATTATCAATAACAGGAGATTCGCGATGAT 1140
Db 930 CTGTTGCCCTTCCAGTGGCAGATCAATTATCAATAACAGGAGATTCGCGATGAT 989
Qy 1141 GTACTTCTCAGTACATTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 990 GTACTTCTCAGTACATTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049
Qy 1201 CAACCCATCTCTACACCTCATTTCAAGAAGTACAGAGCGCGCGCTTTAACTGCT 1260
Db 1050 CAACCCATCTCTACACCTCATTTCAAGAAGTACAGAGCGCGCGCTTTAACTGCT 1109
Qy 1261 GCTCGAAGGAAGTCCAGGCGGAGAGGCTTCCAGAGGAGGACACTCGGGGGAGT 1320
Db 1110 GCTCGAAGGAAGTCCAGGCGGAGAGGCTTCCAGAGGAGGACACTCGGGGGAGT 1169
Qy 1321 TGCAGGGGACACTGGAGAGACAGCTGGGCTACACGAGACAGCGCTAACGTGAAGAC 1380
Db 1170 TGCAGGGGACACTGGAGAGACAGCTGGGCTACACGAGACAGCGCTAACGTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

RESULT 5
AF034632
LOCUS AF034632 2040 bp DNA linear PRI 13-JUL-1998
DEFINITION Homo sapiens orphan G protein-coupled receptor (GPR38) gene,
complete cds.
ACCESSION AF034632
VERSION AF034632.1 GI:2654158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2040)
AUTHORS McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighner,S.D.,
Hreniuk,D.L., Smith,R.G., Howard,A.D. and Van der Ploeg,L.H.
TITLE Cloning and characterization of two human G protein-coupled
receptor genes (GPR38 and GPR39) related to the growth hormone
secretagogue and neurotensin receptors
JOURNAL Genomics 46 (3), 426-434 (1997)
MEDLINE 98110578
PUBMED 9441746
REFERENCE 2 (bases 1 to 2040)
AUTHORS McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighner,S.D.,
Hreniuk,D.L., Smith,R.G., Van Der Ploeg,L.H.T. and Howard,A.D.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co.,
Inc., PO Box 2000, Rahway, NJ 07065, USA
FEATURES
source
1..2040
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="13q14-21"
..2040
/genes="GPR38"
join(1..901,1703..2040)
/genes="GPR38"
/notes="orphan G protein-coupled receptor related to growth
hormone secretagogue and neurotensin receptor"
/codon_start=1
/product="GPR38"
/protein_id="AAC26081.1"
/db_xref="GI:2654159"
/translation="MGSPNGSDGPEGAREPWPALPPCDERCSPPPLGALVPTAV
CICLPVGGNVLRIYRDMRTTNLYGSMVADLLILGLPFDLYRLMRSR
PMWFGPLRLSLVYGGCTYATLLHMTALSVERYLAICRPLRLVLTNRVRLAIA
VLMAVALSAGPFLFLVGVQDPIISVPLNGTARIASSPLASSPPLMSRAPPS
PSGPTAEAAALPSRCPSPALQALRVMLVTTAYFPFLFLSLGLIGLELWS
SRPLRGPASGRGRHROTIVLLVVLAFIICWLPFHVGRITVINTEDSRMYSQ
YFNIVALQIPLYLSASINPLIYNLISKYRAAPFLKLLARKSRPGRHRSRDAGEVAG
DTGDTVGTGTTSTANVTMG"
BASE COUNT 347 a 631 c 598 g 464 t
ORIGIN
Query Match 75.7%; Score 1052; DB 9; Length 2040;
Best Local Similarity 100.0%; Pred. No. 1.1e-130; Indels 0; Gaps 0;
Matches 1052; Conservative 0; Mismatches 0; Gaps 0;
Qy 1 ATGGGCGAGCCCTGGAAACCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 1 ATGGGCGAGCCCTGGAAACCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Qy 61 GCCTGCGCGCTTTCGACGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 GCCTGCGCGCTTTCGACGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Qy 121 GTACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACCAACCAACCAACCAACCAACCAATG 240
Db 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACCAACCAACCAACCAACCAACCAATG 240
Qy 241 GCCGTGTCCGACCTACTCATCTGCTGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCGC 300
Db 241 GCCGTGTCCGACCTACTCATCTGCTGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCGC 300
Qy 301 TCCGCGCCCTGGGTGTTTCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCCGCGCCCTGGGTGTTTCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Qy 361 TGCACCTAGCCACGCTGTCACATGACCGCGCTCAGCGTCGAGCGCTACCTGGCCATC 420
 Db 361 TGCACCTAGCCACGCTGTCACATGACCGCGCTCAGCGTCGAGCGCTACCTGGCCATC 420
 Qy 421 TGGCGCCGCTCCCGCGCCCGCTTGGTACACCGCGCGCGCTCCGCGGCTCATCGCT 480
 Db 421 TGGCGCCGCTCCCGCGCCCGCTTGGTACACCGCGCGCGCTCCGCGGCTCATCGCT 480
 Qy 481 GTGCTCTGGGCGCTGGCGCTCTCTCTCGCGGTCCTTCTTGTCTGGTGGGCGTCGAG 540
 Db 481 GTGCTCTGGGCGCTGGCGCTCTCTCTCGCGGTCCTTCTTGTCTGGTGGGCGTCGAG 540
 Qy 541 CAGGACCCGCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCTCG 600
 Db 541 CAGGACCCGCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCTCG 600
 Qy 601 CCTCTCGCTCTGTCGCGCTCTCTCTGCTCTGCGGGCGGCGGCGGCTCCGCGCTCG 660
 Db 601 CCTCTCGCTCTGTCGCGCTCTCTCTGCTCTGCGGGCGGCGGCGGCTCCGCGCTCG 660
 Qy 661 GGGCCCGAGACCGCGGAGCGCGGCTGTTACAGCGCGAATGCGCGCGGCGGCGCGG 720
 Db 661 GGGCCCGAGACCGCGGAGCGCGGCTGTTACAGCGCGAATGCGCGCGGCGGCGCGG 720
 Qy 721 CAGCTGGGCGCGCTGCTGTCATGCTGTGGGTACACCGCGCTACTTCTTCTGCGCTTT 780
 Db 721 CAGCTGGGCGCGCTGCTGTCATGCTGTGGGTACACCGCGCTACTTCTTCTGCGCTTT 780
 Qy 781 CTGTGCTCAGATCTCTTACGGGCTCATCGCGGCGGAGCTGTGGAGCAGCGCGGCGG 840
 Db 781 CTGTGCTCAGATCTCTTACGGGCTCATCGCGGCGGAGCTGTGGAGCAGCGCGGCGG 840
 Qy 841 CTGGAGCGCGCGCTCGGGCGGAGAGGCGCACCGGAGACCGCTCCGCGTCTG 900
 Db 841 CTGGAGCGCGCGCTCGGGCGGAGAGGCGCACCGGAGACCGCTCCGCGTCTG 900
 Qy 901 CGTAAGTGGAGCGCGCTGTTCAAGAGACCGCTGCTCAGTCCCGCGCGCGGAGCC 960
 Db 901 CGTAAGTGGAGCGCGCTGTTCAAGAGACCGCTGCTCAGTCCCGCGCGCGGAGCC 960
 Qy 961 GCGCAACGCTGGTCCCTTCCCTGCTCGCCAGCTCTGGGGCGCGCTTCAGTCC 1020
 Db 961 GCGCAACGCTGGTCCCTTCCCTGCTCGCCAGCTCTGGGGCGCGCTTCAGTCC 1020
 Qy 1021 TTTCCTATTTTCAGTTCAGCTCCACCGCGG 1052
 Db 1021 TTTCCTATTTTCAGTTCAGCTCCACCGCGG 1052

RESULT 6
 AL137000
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-203116 on chromosome 13 Contains the gene for KIAA0970 protein, COX7CP1 (cytochrome c oxidase subunit VIIC pseudogene 1), a novel pseudogene, the GPR38 (G protein-coupled receptor 38) gene, ESTs, STSS, GSSs and a CpG island, complete sequence.
 AL137000
 VERSION AL137000.6 GI:9944121
 KEYWORDS HTG; COX7CP1; CpG island; cytochrome c oxidase; G protein-coupled receptor; GPR38; KIAA0970.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 163284)
 Wall, M.
 Direct Submission
 Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 JOURNAL requests: Clonerequests@sanger.ac.uk
 On Aug 29, 2000 this sequence version replaced gi:9926419.
 COMMENT During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>
 IMPORTANT: This sequence is not the entire insert of clone RP11-203116. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
 The true left end of clone RP11-452110 is at 1 in this sequence. The true right end of clone RP11-103118 is at 163284 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-203116 is from the library RPII-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6.

FEATURES
 source
 1. .163284
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-203116"
 /clone_lib="RPII-11.1"
 3. .121
 /note="MER20 repeat: matches 4. .126 of consensus"
 152. .261
 /note="MIR repeat: matches 20. .137 of consensus"
 262. .554
 /note="AluX repeat: matches 1. .293 of consensus"
 555. .698
 /note="MIR repeat: matches 137. .255 of consensus"
 1012. .1291
 /note="AluJo repeat: matches 7. .291 of consensus"
 1320. .1425
 /note="MER94 repeat: matches 4. .114 of consensus"
 1591. .1804
 /note="L1PAB repeat: matches 5945. .6158 of consensus"
 3240. .3621
 /note="L2 repeat: matches 2310. .2691 of consensus"
 3824. .4106
 /note="AluJo repeat: matches 1. .301 of consensus"
 4975. .5042
 /note="L2 repeat: matches 2639. .2705 of consensus"
 6113. .6408
 /note="AluX repeat: matches 1. .295 of consensus"
 7646. .7732
 /note="AluSc repeat: matches 1. .85 of consensus"
 9887. .10191
 /note="AluX repeat: matches 1. .306 of consensus"
 12201. .12247
 /note="L2 repeat: matches 2704. .2750 of consensus"
 12472. .12738
 /note="L2 repeat: matches 1304. .1530 of consensus"
 12739. .12874
 /note="FLAM_C repeat: matches 2. .133 of consensus"
 12875. .12887
 /note="L2 repeat: matches 1530. .1542 of consensus"
 12922. .13227
 /note="L1MB3 repeat: matches 5871. .6183 of consensus"
 13303. .13348


```
Qy 541 CAGACACCCGGCATCTCCGTAGTCCCGGGCTCAATGCGACCGCGCGGATCGCTCCCTCG 600
Db 144888 CAGACACCCGGCATCTCCGTAGTCCCGGGCTCAATGCGACCGCGCGGATCGCTCCCTCG 144947
Qy 601 CTTCTCGCTCGTCCCGCTCTCTGCTCTCGGGGGGCGACCGCGCTCCCGCGCTCG 660
Db 144948 CTTCTCGCTCGTCCCGCTCTCTGCTCTCGGGGGGCGACCGCGCTCCCGCGCTCG 145007
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACCGCGGAAATGCGCGCGGAGCGCGCGG 720
Db 145008 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACCGCGGAAATGCGCGCGGAGCGCGCGG 145067
Qy 721 CAGCTGGCGCGCTCGCTGATCTGCTGATCTGAGTACACCGCTCTCTCTCTCTCTCT 780
Db 145068 CAGCTGGCGCGCTCGCTGATCTGCTGATCTGAGTACACCGCTCTCTCTCTCTCTCT 145127
Qy 781 CTGTGCTCAGATCTCTACGGGCTCATCGGGGAGAGCTGTGAGACAGCGCGCGCGG 840
Db 145128 CTGTGCTCAGATCTCTACGGGCTCATCGGGGAGAGCTGTGAGACAGCGCGCGCGG 145187
Qy 841 CTGGAGGCGCGCGCTCGGGGAGAGAGGCGACCGCGAGACCGTCCGCTCTCTG 900
Db 145188 CTGGAGGCGCGCGCTCGGGGAGAGAGGCGACCGCGAGACCGTCCGCTCTCTG 145247
Qy 901 CGTAAGTGGAGCGCGCTGTTCCAAAGACCGCTCGCTGAGTCCCGCGCGCGGAGC 960
Db 145248 CGTAAGTGGAGCGCGCTGTTCCAAAGACCGCTCGCTGAGTCCCGCGCGGAGC 145307
Qy 961 GCGCAACCGTGGTCCCTTCCCTGCTGCTGCGCGAGCTGTGGGCGCGCTTCCAGTCCC 1020
Db 145308 GCGCAACCGTGGTCCCTTCCCTGCTGCTGCGCGAGCTGTGGGCGCGCTTCCAGTCCC 145367
Qy 1021 TTTCTATTTCGATTCAGCTCCACCGCGG 1052
Db 145368 TTTCTATTTCGATTCAGCTCCACCGCGG 145399

RESULT 7
AL596304
LOCUS Homo sapiens chromosome 13 clone RP11-804, *** SEQUENCING IN
DEFINITION AL596304
ACCESSION AL596304
VERSION AL596304.3 GI:15147703
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Clark G.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Aug 10, 2001 this sequence version replaced gi:15131387.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA804
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 149716 bases at least Q40
Consensus quality: 150039 bases at least Q30
Consensus quality: 150247 bases at least Q20
Insert size: 341360; sum-of-ontigs
Insert size: 157238; 2.0% error; agarose-fp
Quality coverage: 3.92x in Q20 bases; sum-of-contigs Quality
```

```
coverage: 8.77x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
* 1 146853: contig of 146853 bp in length
* 146854 146953: gap of 100 bp
* 146954 338660: contig of 191707 bp in length
* 338661 338760: gap of 100 bp
* 338761 341560: contig of 2800 bp in length.
FEATURES
Location/Qualifiers
source
1..341560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="13"
/clone="RP11-804"
/clone_lib="RPC1-11.1"
misc_feature
1..146853
/notes="assembly_fragment:03377.0"
misc_feature
146954..338660
/notes="assembly_fragment:03377.1"
misc_feature
338761..341560
/notes="assembly_fragment:03476"
BASE COUNT 101629 a 64391 c 66276 g 109064 t 200 others
ORIGIN
Query Match 75.7%; Score 1052; DB 2; Length 341560;
Best Local Similarity 100.0%; Pred. No. 3.1e-131;
Matches 1052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGAGCCCTTGGAAACGCGACGCGCCCGCGGCGCGCGCGCGCGCGCGCGCGCG 60
Db 319724 ATGGCGAGCCCTTGGAAACGCGACGCGCCCGCGGCGCGCGCGCGCGCGCGCG 319783
Qy 61 GCGCTCGCGCTTGGCGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGCTGGTCCG 120
Db 319784 GCGCTCGCGCTTGGCGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGCTGGTCCG 319843
Qy 121 GTGACCGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 319844 GTGACCGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319903
Qy 181 ATGCTGATCGGGCGCTTACCGGGACATGCGGACCAACCAACCAACCAACCAACCA 240
Db 319904 ATGCTGATCGGGCGCTTACCGGGACATGCGGACCAACCAACCAACCAACCAAC 319963
Qy 241 GCGGTGTGCGACCTACTATCTGCTGCGGGTCCGCTGCGCTGCGCTGCGCTGCGCTG 300
Db 319964 GCGGTGTGCGACCTACTATCTGCTGCGGGTCCGCTGCGCTGCGCTGCGCTGCGCTG 320023
Qy 301 TCGCGCGCTTGGGTGTTGGGGCGCGCTGCTGCGCGCTGCTGCGCTGCTGCGCTGCT 360
Db 320024 TCGCGCGCTTGGGTGTTGGGGCGCGCTGCTGCGCGCTGCTGCGCTGCTGCGCTG 320083
Qy 361 TGCACCTAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 320084 TGCACCTAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 320143
Qy 421 TGCAGCCCGCTCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 320144 TGCAGCCCGCTCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 320203
Qy 481 GTGCTCTGGGCGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 320204 GTGCTCTGGGCGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320263
Qy 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGCGACCGCGCGGATCGCTCCCTCG 600
```

320264	Db		CAGAACCCCGCATCTCGTATGCCGGCCTCAATGACACCGCGGATCGCTCTCTCG	320323
601	Qy		CCTCTCGCCTCGTCCGCCCTCTCTGGCTCTCGCGGGCGCACCGCCGTCCCGCGCTCG	660
320324	Db		CCTCTCGCCTCGTGGCGCCTCTCTGGCTCTCGCGGGCGCACCGCGTCCCGCGCTCG	320383
661	Qy		GGGCCCCGAGACCGCGGAGCGCGCGCTGTTCACGCGCGAATGCGGCGAGCCCCGCG	720
320384	Db		GGGCCCCGAGACCGCGGAGCGCGCGCTGTTCAGCGCGAATGCGGCGAGCCCCGCG	320443
721	Qy		CAGCTGGGCGCGTGGTGTCATGCTGGTGGGTACACCGCCTACTTCTTCCTGGCCCTTT	780
320444	Db		CAGCTGGGCGCGTGGTGTCATGCTGGTGGGTACACCGCCTACTTCTTCCTGGCCCTTT	320503
781	Qy		CTGTGCCTCAGCATCCTCTACGGGTCATCGGGCGGGAGCTGTGAGACAGCGCGGCGCG	840
320504	Db		CTGTGCCTCAGCATCCTCTACGGGTCATCGGGCGGGAGCTGTGAGACAGCGCGGCGCG	320563
841	Qy		CTGCGAGGCCGCGCGCTCTCGGGCGGGAGAGAGCCACCGGACACGCTCGCGTCTCTG	900
320564	Db		CTGCGAGGCCGCGCGCTCTCGGGCGGGAGAGAGAGCCACCGGACACGCTCGCGTCTCTG	320623
901	Qy		CGTAAGTGGAGCGCGGTGGTTCAAAGACGCTGCCTGCAGTCGCGCCCGCGGGGACC	960
320624	Db		CGTAAGTGGAGCGCGGTGGTTCAAAGACGCTGCCTGCAGTCGCGCCCGCGGGGACC	320683
961	Qy		GCGCAAAAGCTGGGTCCCTTCCCTGCTCGGCCAGCTCTGGGGCGCGCTTCAGCTCC	1020
320684	Db		GCGCAAAAGCTGGGTCCCTTCCCTGCTCGGCCAGCTCTGGGGCGCGCTTCAGCTCC	320743
1021	Qy		TTTCCTATTTCGATTCCAGCTCCACCCGCG	1052
320744	Db		TTTCCTATTTCGATTCCAGCTCCACCCGCG	320775

RESULT 8
 AX711879
 LOCUS
 DEFINITION
 Accession
 AX711879
 VERSION
 AX711879.1
 KEYWORDS
 GI:29787684
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1
 AUTHORS
 Zhang, Y., Moffatt, M., Cookson, W. and Tinsley, J.O.
 TITLE
 Copy
 JOURNAL
 Patent: WO 03000727-A 1 03-JAN-2003;
 ISIS INNOVATION LIMITED (GB)
 FEATURES
 Location/Qualifiers
 1..349980
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="Original length of seq 1 to big: 410,846-Splitted
 into-seq 1: from 000,001 to 349,980-seq 327: from 300,001
 to 410,846-Original length of seq 71 to big:
 379,652-Splitted into-seq 71: from 000,001 to 349,980-seq
 328: from 300,001 to 379,652"
 BASE COUNT 106097 a 69968 c 69717 g 105198 t
 ORIGIN
 Query Match 75.7%; Score 1052; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 3.1e-11;
 Matches 1052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGGGCAGCCCTGTGAACGGCAGCGCCGCCGAGGGGGCCGGAGAGCCGCGTGGCCCC 60
 Db 38016 ATGGGCAGCCCTGTGAACGGCAGCGCCGCCGAGGGGGCCGGAGAGCCGCGTGGCCCC 38075

RESULT 9
AX739961
LOCUS
DEFINITION


```
Db 80 GCGTGGCGGCTTGGACGAGCGCGTCTGCGCCCTTTCCCTTGGGGGCGCTGGTCCG 139
Qy 121 GTGACCGCTGTGTCCTGTGCTGTCTGCTCGTGGGGTGAGCGCAACGTGGTACCGTG 180
Db 140 GTGACCGCTGTGTCCTGTGCTGTCTGCTCGTGGGGTGAGCGCAACGTGGTACCGTG 199
Qy 181 ATGCTGATCGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATG 240
Db 200 ATGCTGATCGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATG 259
Qy 241 GCCGTGTCGACCTACTCATCTGCTCGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 300
Db 260 GCCGTGTCGACCTACTCATCTGCTCGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 319
Qy 301 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 360
Db 320 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 379
Qy 361 TGCACCTAGCCACCGCTGCTGACATGACCGCGCTGACGCGTGGCGTGGCGTGGCGTGGCG 420
Db 380 TGCACCTAGCCACCGCTGCTGACATGACCGCGCTGACGCGTGGCGTGGCGTGGCGTGGCG 439
Qy 421 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 540
Db 440 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 499
Qy 481 GTGCTGTGGGGCGTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 540
Db 500 GTGCTGTGGGGCGTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 559
Qy 541 CAGGACCGCGCATCTCGTGTGTCGCGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 600
Db 560 CAGGACCGCGCATCTCGTGTGTCGCGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 619
Qy 601 CCTCTGCGCTGTCGCGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 660
Db 620 CTTTGTGCTGTCGCGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 674
Qy 661 GGGCGCGAGACCGCG 675
Db 675 GGGCGCGAGACCGCG 689
```

```
RESULT 11
HSA342408 692 bp DNA linear PRI 18-JUL-2002
LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone
DEFINITION NRI-PB24C.
VERSION AJ342408
KEYWORDS AJ342408.1 GI:15886883
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 692)
Kutsenko A.S., Gizatullin R.Z., Al-Amin A.N., Wang F., Kvasha S.M.,
Podowski R.M., Matushkin Y.G., Gyanchandani A., Muravenko O.V.,
Levitky V.G., Kolchanov A.I., Protopopov A.I., Kashuba V.I.,
Kiselev L.L., Wasserman W., Wahlstedt C. and Zabarovsky E.R.
Not1 flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
JOURNAL 2131767
MEDLINE 12136098
PUBMED
REFERENCE 2 (bases 1 to 692)
AUTHORS Zabarovsky E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
Location/Qualifiers
```

```
source 1. .692
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NRI-PB24C"
BASE COUNT 92 a 239 c 219 g 140 t 2 others
ORIGIN
Query Match 41.2%; Score 572.2; DB 9; Length 692;
Best Local Similarity 91.5%; Pred. No. 7.9e-67;
Matches 616; Conservative 0; Mismatches 55; Indels 2; Gaps 1;
Qy 1 ATGGGAGAGCCCTTGGAAACGACGACGCGCCCGGAGGGGCGGAGGAGCGCCGTCGCC 60
Db 20 ATGGGAGAGCCCTTGGAAACGACGACGCGCCCGGAGGGGCGGAGGAGCGCCGTCGCC 79
Qy 61 GCCTGCGCGCTTGGACGAGCGCGCTGCTGCGCCCTTTCCCTTGGGGGCGCTGGTCCG 120
Db 80 GCCTGCGCGCTTGGACGAGCGCGCTGCTGCGCCCTTTCCCTTGGGGGCGCTGGTCCG 139
Qy 121 GTGACCGCTGTGTCCTGTGCTGTGCTGCTGCGGCGTGGCGCAACGTGGTACCGTG 180
Db 140 GTGACCGCTGTGTCCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199
Qy 181 ATGCTGATCGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATG 240
Db 200 ATGCTGATCGGGCGCTACCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATCGGGAATG 259
Qy 241 GCCGTGTCGACCTACTCATCTGCTCGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 300
Db 260 GCCGTGTCGACCTACTCATCTGCTCGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 319
Qy 301 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 360
Db 320 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 379
Qy 361 TGACCTAGCCACCGCTGCTGACATGACCGCGCTGACGCGTGGCGTGGCGTGGCGTGGCG 420
Db 380 TGACCTAGCCACCGCTGCTGACATGACCGCGCTGACGCGTGGCGTGGCGTGGCGTGGCG 439
Qy 421 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 480
Db 440 TCGCGGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 499
Qy 481 GTGCTGTGGGGCGTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 540
Db 500 GTGCTGTGGGGCGTGGGGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 559
Qy 541 CAGGACCGCGCATCTCGTGTGTCGCGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 600
Db 560 CATGACACCTGTATCTCCGTAGTCTCGGGCTCAATGCCACCGCGGATAGCTACTTC 619
Qy 601 CCT--CTGCGCTGTCGCGCGCTGCTGCGCGCGTGTCCCTTACGTGGGGGAGGGC 658
Db 620 GCTTATTAGCATAGTCGACGACTCTGTTGATATCGGGGTGCAACCGAGCATACATACCGT 679
Qy 659 CGGGGCGCGGAGAC 671
Db 680 CAGGGCGCGATAC 692
RESULT 12
HSA339459
LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone
DEFINITION NRI-WC14C.
VERSION AJ339459
KEYWORDS AJ339459.1 GI:15883877
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
```

REFERENCE 1 (bases 1 to 615)
AUTHORS Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M., Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V., Levitsky,V.G., Kolchanov,N.A., Protopopov,A.I., Kashuba,V.I., Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
TITLE NotI flanking sequences: a tool for gene discovery and verification of the human genome
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
MEDLINE 22131767
PUBMED 12136098
REFERENCE 2 (bases 1 to 615)
AUTHORS Zabarovsky,E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden

FEATURES
source
1. .615
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NR1-WC14C"
BASE COUNT 61 a 224 c 204 g 121 t 5 others
ORIGIN
Query Match 40.1%; Score 557.4; DB 9; Length 615;
Best Local Similarity 95.6%; Pred. No. 7.6e-65;
Matches 570; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 ATGGGAGCCCTGGAAACGGAGCGACGGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60
Db 20 ATGGGAGCCCTGGAAACGGAGCGACGGCCCGAGGGGGCGGGAGCGCCGTGGCCC 79
Qy 61 GCGCTGCGCGCTTGGAGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGCTGGTGGC 120
Db 80 GCGCTGCGCGCTTGGAGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGCTGGTGGC 139
Qy 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 140 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 199
Qy 181 ATGCTGATCGGGGCGCTACCGGAGCATCGGAGACACCAACCTGTGACCTGGGAGCATG 240
Db 200 ATGCTGATCGGGGCGCTACCGGAGCATCGGAGACACCAACCTGTGACCTGGGAGCATG 259
Qy 241 GCGGTGTCGACCTACTCATCTGCTGGGCTGCGGTTCGACCTGTACCGCTCTGGCGC 300
Db 260 GCGGTGTCGACCTACTCATCTGCTGGGCTGCGGTTCGACCTGTACCGCTCTGGCGC 319
Qy 301 TCGCGGCGCTTGGGTGTGGGGCGCGCTGCTTGGCGCGCTGTCCCTCTACGTGGGCGAGGC 360
Db 320 TCGCGGCGCTTGGGTGTGGGGCGCGCTGCTTGGCGCGCTGTCCCTCTACGTGGGCGAGGC 379
Qy 361 TGCACTTACGCCACCGCTGTGCAATGACCGCGCTAGCGTCGAGCGCTACCTGGCCATC 420
Db 380 TGCACTTACGCCACCGCTGTGCAATGACCGCGCTAGCGTCGAGCGCTACCTGGCCATN 439
Qy 421 TGCGCGCGCTTGGGCGCGCGCTTGGTCAACCGGCGCGCGCTCGCGCGCTCATCGCT 480
Db 440 TGCNCGCGCTTGGGCGCGCGCTTGGGTCAACCGGCGCGCGCTCGCGCGCTCATCGCT 499
Qy 481 GTGCTTGGGCGCGTGGGCGCTGCTCTTGGCGCGCTTCTTGTTCCTGGTGGGCGTGGAG 540
Db 500 GTGCTTGGGCGCGTGGGCGCTTGTCTTGGCGCGCTTCTTGTTCCTGGTGGGCGTNNAA 559
Qy 541 CAGAACCCCGCATCTCCGTAGTCCGGGCTCAATGGCACCGCGGATCGGCTC 596
Db 560 CAAGAACCCCGCATTTCCGTAGTCCGGGCTTAAATGGCACCGCGGATATGCTC 615

RESULT 13
HSA326768. 701 bp DNA linear PRI 18-JUL-2002
LOCUS HSA326768

DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone NR5-IP16C.
ACCESSION AJ326768
VERSION AJ326768.1 GI:158711186
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 701)
AUTHORS Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M., Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V., Levitsky,V.G., Kolchanov,N.A., Protopopov,A.I., Kashuba,V.I., Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
TITLE NotI flanking sequences: a tool for gene discovery and verification of the human genome
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
MEDLINE 22131767
PUBMED 12136098
REFERENCE 2 (bases 1 to 701)
AUTHORS Zabarovsky,E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden

FEATURES
source
1. .701
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NR5-IP16C"
BASE COUNT 108 a 218 c 235 g 139 t 1 others
ORIGIN
Query Match 30.8%; Score 424.8; DB 9; Length 701;
Best Local Similarity 81.1%; Pred. No. 3e-47;
Matches 554; Conservative 0; Mismatches 123; Indels 6; Gaps 5;
Qy 1 ATGGGAGCCCTGGAAACGGAGCGACGGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60
Db 20 ATGGGAGCCCTGGAAACGGAGCGACGGCCCGAGGGGGCGGGAGCGCCGTGGCCC 79
Qy 61 GCGCTGCGCGCTTGGAGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGCTGGTGGC 120
Db 80 GCGCTACCGACATGCTACGATCGCGCTGCTCTCCCTTCCCTGGGGGCGCTGGTGGC 139
Qy 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 140 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 199
Qy 181 ATGCTGATCGGGGCGCTACCGGAGCATCGGAGACACCAACCTGTGACCTGGGAGCAT 239
Db 200 ATGATAACCGGCGCTCCCGGAGCATGCGGAGACACCAACCTGTGACCTGGGAGCAT 259
Qy 240 GGCCTGTGCTCGGCTACTCATCTGCTCGGGCTGCGCTTTCGACCTGTACCGCTCTGGCG 299
Db 260 GGACAGGTTCACATATCTTATCTGCTCGGGCTGCGCTTAGACCTAGACACGCTCTGGCG 319
Qy 300 CTGCGCGCGCTGGGTGTGCGGCGCGCTGCTGTGCGCGCTGTCTCTACGTGGGCGAGGG 359
Db 320 CTGCGGAGCTTGGGAGAGCGGCGGTGCTGTGACGCTCTCTCTACGTGGGCGAGGG 379
Qy 360 CTGCACTACGCGCGCTGCTGCATGACCGGCTCAGGCTCAGGCTACCTGGCCAT 419
Db 380 CTGCACTTCTCCACGCTGTGGACATGACCGGCTCAGGCTAGAGCGCTACCTGGGCGAT 439
Qy 420 CTGCGCGCGCGCTCCGCGCGCGCTCTTGTGTACCCGCGCGCGCTCCGCGCGCTCATCGC 479
Db 440 TTGCCACCGATCCGACCCCGCGTTTGTGTACCCAGAGCCGAGTTCGCGCGCTTATCTC 499
Qy 480 TGTGCTCTGGGCGCGTGGCGCTGCTCTCTGCGGCTTCTTGTGCTCTGTGCGGCTCGA 539
Db 500 TGTGCTCTGGGCGAGGTGCTGGTCTATGCGCGTACCTTATTGCTCTTGTGGGGGTT--A 557

Qy 540 GCAGGACCCCGGATCTCCGTAGTCCGGCGCTCAATGGGACC-GCCGGGATCGCTCCT 598
 |||||
 Db 558 AGCAGACCCCGGATATGGTAGTACCGGGCTTAATGGCACCAGCGGATCGTCTCCT 617
 |||||
 Qy 599 CGCTCTCGCTCGT-CGCCGCTCTCTGGCTCCGGGGCGCC-ACCGCGTCCCGGCC 656
 |||||
 Db 618 GGAATATAGTTTGTGGCCACATGTTGGTAGTGGCGGGCCACCGACGCTCCACAC 677
 |||||
 Qy 657 GTGGGGCCCGAGACCGCGGAGG 679
 |||||
 Db 678 GGTGTGCACACAGACCGCGGAGG 700
 |||||

RESULT 14
 AC109158/c
 LOCUS
 DEFINITION Mus musculus clone RP24-199E5, WORKING DRAFT SEQUENCE, 16 ordered pieces.

AC109158
 AC109158.3 GI:31416079
 HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Birren,B., Nussbaum,C. and Lander,E.
 1 (bases 1 to 192116)

Unpublished
 Mus musculus, clone RP24-199E5

2 (bases 1 to 192116)

Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galegan,J., Gardyna,S., Glnde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,C., Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 192116)

Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B., DeArallano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galegan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,

Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (05-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 5, 2003 this sequence version replaced gi:20303716.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu

----- Project Information

Center project name: L18656

Center clone name: 199_E_5

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 188041 bases at least Q40

Consensus quality: 189224 bases at least Q30

Consensus quality: 189863 bases at least Q20

Insert size: 176000; agarose-fp

Insert size: 190616; sum-of-contigs

Quality coverage: 11.9 in Q20 bases; agarose-fp

Quality coverage: 11.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 16 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 291: contig of 291 bp in length

* 292 391: gap of 100 bp

* 392 1310: contig of 919 bp in length

* 1311 1410: gap of 100 bp

* 1411 2084: contig of 674 bp in length

* 2085 2184: gap of 100 bp

* 2185 3934: contig of 1750 bp in length

* 3935 4034: gap of 100 bp

* 4035 7860: contig of 3826 bp in length

* 7861 7960: gap of 100 bp

* 7961 10973: contig of 3013 bp in length

* 10974 11073: gap of 100 bp

* 11074 13977: contig of 2904 bp in length

* 13978 14077: gap of 100 bp

* 14078 26266: contig of 12189 bp in length

* 26267 26366: gap of 100 bp

* 26367 36455: contig of 10089 bp in length

* 36456 47992: contig of 11437 bp in length

* 47993 48092: gap of 100 bp

* 48093 66740: contig of 18647 bp in length

* 66740 66839: gap of 100 bp

* 66839 88335: contig of 21496 bp in length

* 88336 88435: gap of 100 bp

* 88436 124376: contig of 35941 bp in length

* 124377 124476: gap of 100 bp

* 124477 145657: contig of 21181 bp in length

* 145658 145757: gap of 100 bp

* 145758 165431: contig of 19674 bp in length

* 165432 165531: gap of 100 bp

* 165532 192116: contig of 26585 bp in length.

Location/Qualifiers
 1. .192116

FEATURES
 source

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 13:48:11 ; Search time 426.676 Seconds
(without alignments)
8794.080 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atggcagcccttggaacgg.....acgtgaagcgtgggataa 1390

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1390	100.0	1390	21	AAZ45404
2	1161	83.5	1161	22	AAF83684
3	1078	77.6	1239	21	AAZ45403
4	1078	77.6	1239	22	AAF83683
5	1078	77.6	1239	22	AAF85449
6	1078	77.6	1239	24	ABK90132
7	1078	77.6	1239	25	AB242842
8	1078	77.6	1506	22	AA166989

9	1047.2	75.3	2040	21	AAA46116	Human G protein co
10	1038	74.7	3066	21	AAZ45402	Genomic sequence o
11	754.6	54.3	1203	22	AAF85448	Nucleotide sequenc
12	546	39.3	813	22	AAF85447	Nucleotide sequenc
13	501.2	36.1	1179	24	ABQ47146	Oligonucleotide fo
14	501.2	36.1	1179	24	ABQ47147	Oligonucleotide fo
15	437.2	31.5	1179	24	ABQ47148	Oligonucleotide fo
16	437.2	31.5	1179	24	ABQ47149	Oligonucleotide fo
17	283	20.4	283	19	AAV44930	Galanin receptor G
18	283	20.4	283	19	AAV32651	Galanin receptor G
19	283	20.4	283	19	AAV28290	Galanin receptor G
20	283	20.4	283	24	ABK14060	Rat galanin recept
21	248.2	17.9	1850	21	AAZ61492	cDNA encoding canl
22	234.8	16.9	1063	18	AAZ68662	Pig growth hormone
23	234.8	16.9	1063	18	AAZ69754	Swine growth hormo
24	233.2	16.8	1029	18	AAZ68663	Pig growth hormone
25	233.2	16.8	1029	18	AAZ69755	Swine growth hormo
26	231.6	16.7	1095	21	AAZ45993	cDNA encoding the
27	231.6	16.7	4009	21	AAZ45967	DNA encoding the m
28	229	16.5	250	25	ACA55761	Pig signalling pat
29	228.8	16.5	870	25	ABZ42674	Human growth hormo
30	228.8	16.5	1122	18	AAZ68665	Human growth hormo
31	228.4	16.4	1092	22	AAZ27800	Rat growth hormone
32	228.4	16.4	3129	18	AAZ68667	Rat growth hormone
33	228.4	16.4	3129	18	AAZ69759	Rat growth hormone
34	227.8	16.4	1088	18	AAZ68664	Human growth hormo
35	227.8	16.4	1088	18	AAZ69756	Human growth hormo
36	227.8	16.4	1101	21	AAZ30643	Human G protein-co
37	227.8	16.4	1101	21	AAZ30732	DNA encoding human
38	227.8	16.4	1101	21	AAZ51463	Human G protein-co
39	227.8	16.4	1101	22	AAZ83680	Human G-protein co
40	227.8	16.4	1101	24	AAZ30395	Human growth-hormo
41	227.8	16.4	1101	25	ACA56550	Human signalling p
42	227.2	16.3	1122	18	AAZ69757	Human growth hormo
43	226.8	16.3	1092	18	AAZ69760	Rat growth hormone
44	225.2	16.2	1092	21	AAZ45405	DNA encoding the p
45	225.2	16.2	1092	22	AAF85450	Nucleotide sequenc

ALIGNMENTS

RESULT 1
AAZ45404
ID AAZ45404 standard; cDNA; 1390 BP.

XX AAZ45404;

AC AAZ45404;

XX 27-MAR-2000 (first entry)

DE CDNA encoding the motilin receptor splice variant MTL-R1B.

XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1a; MTL-R1b; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW endoscopy; duodenal intubation; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 1.1161
FT CDS /tag= a
FT /product= "MTL-R1B"

XX WO9964436-A1.

XX 16-DEC-1999.

PD

RESULT 3
AAZ45403
ID AAZ45403 standard; cDNA; 1239 BP.
XX AC AAZ45403;
DT 27-MAR-2000 (first entry)
XX DE cDNA encoding the motilin receptor splice variant MTL-R1A.
XX KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW endoscopy; duodenal intubation; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..1239
XX FT /*tag= a "MTL-R1A"
XX FT /product= "MTL-R1A"
XX PN WO9964436-A1.
XX PD 16-DEC-1999.
XX PF 08-JUN-1999; 99WO-US12773.
XX PR 12-JUN-1998; 98US-0089098.
XX PA (MERI) MERCK & CO INC.
XX PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
XX PI Pong S, Smith RG;
XX DR WPI; 2000-105868/09.
XX DR P-PSDB; AAY54145.
XX PT Novel receptor protein for screening compounds used in treating
XX PT irritable bowel syndrome, constipation and other gastric conditions
XX PS Claim 4; Fig 2; 44pp; English.
XX CC The present sequence encodes splice variant MTL-R1A of the motilin
XX CC receptor. The gene encodes a G-protein coupled receptor, and is
XX CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
XX CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
XX CC functional seven transmembrane domain form, and MTL-R1B is a truncated
XX CC five transmembrane domain. The MTL-R1 proteins are used to identify
XX CC agonists and antagonists which can be used for treating gastric motility
XX CC disorders, functional defects, disorders secondary to neurological
XX CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
XX CC dysmotility, diabetes, infections, stress-related motility disorders,
XX CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
XX CC constipation, chronic idiopathic pseudo obstruction, acute faecal
XX CC impaction, postoperative ileus, gallstones, infantile colic, irritable
XX CC bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
XX CC diarrhoea. They can also be used in the preparation for colonoscopy,
XX CC endoscopy and duodenal intubation.
XX SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;
Query Match 77.6%; Score 1078; DB 21; Length 1239;
Best Local Similarity 89.1%; Pred. No. 2e-186;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Caps 1;

QY 1 ATGGGAGCCCTTGGAAACGAGCAGCGCCCGGAGGGGCGCGGAGCGCGCGTGGCCCC 60
DB |||||
1 ATGGGAGCCCTTGGAAACGAGCAGCGCCCGGAGGGGCGCGGAGCGCGCGTGGCCCC 60
QY 61 GGGCTGGCGCTTGGAGACGAGCGCGCTGCTGGCCCTTTCCCTTGGGGGGGCTGGTGGCG 120
DB |||||
61 GGGCTGGCGCTTGGAGACGAGCGCGCTGCTGGCCCTTTCCCTTGGGGGGGCTGGTGGCG 120
QY 121 GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB |||||
121 GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 ATGCTGATCGGCGCTTACCGGACATGCGGACACCAACCACTTGTACCTGGGCGACATG 240
DB |||||
181 ATGCTGATCGGCGCTTACCGGACATGCGGACACCAACCACTTGTACCTGGGCGACATG 240
QY 241 GCGGTGTCGACCTACTCATCTCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
DB |||||
241 GCGGTGTCGACCTACTCATCTCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
QY 301 TCGCGCCCTGGGTGTTGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCG 360
DB |||||
301 TCGCGCCCTGGGTGTTGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCG 360
QY 361 TGCACCTAGCCACGCTGCTGACATGACCGGCTCAGCGCTCAGCGCTACCTGGGCATC 420
DB |||||
361 TGCACCTAGCCACGCTGCTGACATGACCGGCTCAGCGCTCAGCGCTACCTGGGCATC 420
QY 421 TCGCGCCCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB |||||
421 TCGCGCCCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 GTGCTCTGGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB |||||
481 GTGCTCTGGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CAGGACCCCGCATCTCCGCTAGTCCCGGCTCAATGGCACCGCGGATCGCTCCTCG 600
DB |||||
541 CAGGACCCCGCATCTCCGCTAGTCCCGGCTCAATGGCACCGCGGATCGCTCCTCG 600
QY 601 CTTCTGCTGCTGCGCGCTCTCTGCTGCTCTCGCGGCGCCACCGCGCTCCCGCGCTCG 660
DB |||||
601 CTTCTGCTGCTGCGCGCTCTCTGCTGCTCTCGCGGCGCCACCGCGCTCCCGCGCTCG 660
QY 661 GGGCCGAGACCGCGGAGCGCGGCTGTTGACCGCGGAATGCGCGGAGCGCGCG 720
DB |||||
661 GGGCCGAGACCGCGGAGCGCGGCTGTTGACCGCGGAATGCGCGGAGCGCGCGCG 720
QY 721 CAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 780
DB |||||
721 CAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 780
QY 781 CTGTGCTCAGCATCTCTACCGGCTCATCTGCGGCGGAGCTGTGGAGCAGCGCGCGCG 840
DB |||||
781 CTGTGCTCAGCATCTCTACCGGCTCATCTGCGGCGGAGCTGTGGAGCAGCGCGCGCG 840
QY 841 CTGCGAGCGCGCGCTGCGGCGGAGAGAGGCGCACCGGAGACCGCTGCGGCTGCTG 900
DB |||||
841 CTGCGAGCGCGCGCTGCGGCGGAGAGAGGCGCACCGGAGACCGCTGCGGCTGCTG 900
QY 901 CGTAAGTGAGCGCGCTGTTTCAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB |||||
901 C----- 901
QY 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB |||||
902 ----- 901
QY 1021 TTTCTATTTCGATTCAGCCTCCACCCCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
DB |||||
902 -----TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 929
QY 1081 CTGGTTGCCCTTCCAGCTTGGCAGAATCATTTATACATAACACGGAAGATTTCGGGATGAT 1140

```

Db 930 CTGTTCCCTTCCAGTTCGAGATCAATTTATATAAACACGGAAGATTGCGGATGAT 989
Qy 1141 GTACTTCTCAGTACTTTAAACATCGTCGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
Db 990 GTACTTCTCAGTACTTTAAACATCGTCGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
Qy 1201 CAACCCAAATCTCTACAACTCTATTCAAGAAGTACAGAGCGCGGCTTTAACTGCT 1260
Db 1050 CAACCCAAATCTCTACAACTCTATTCAAGAAGTACAGAGCGCGGCTTTAACTGCT 1109
Qy 1261 GCTCGCAAGAGTCCAGGCGGAGAGGCTTCCACAGAGCAGGACACTCGGGGGAGT 1320
Db 1110 GCTCGCAAGAGTCCAGGCGGAGAGGCTTCCACAGAGCAGGACACTCGGGGGAGT 1169
Qy 1321 TGACGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCCCTAACTGAAGAC 1380
Db 1170 TGACGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCCCTAACTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

```

RESULT 4

AAF83683 standard; DNA; 1239 BP.

AAF83683;

23-JUL-2001 (first entry)

Long form of motilin receptor, GPR-38A isoform encoding DNA.

zslg33; signal transduction; hormone; enzyme; neural development;
gastric contractility; nutrient uptake; digestive; pancreatic; human;
insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R; de;
G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.

Homo sapiens.

Key Location/Qualifiers
CDS 1..1239
/*tag= a
/product= "GPR-38A"

W0200138355-A2.

31-MAY-2001.

22-NOV-2000; 2000WO-US32074.

22-NOV-1999; 99US-0166765.

(ZYMO) ZYMOGENETICS INC.

Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;

WPI; 2001-355879/37.

P-PSDB; AAB62652.

Forming reversible peptide receptor complex for purifying cell and
peptides, stimulating signal transduction and modulating hormone
secretion, involves contacting a receptor with zslg33 polypeptide

Disclosure; Page 102-104; 11pp; English.

The invention relates to a method of forming a reversible peptide-
receptor complex that involves providing an immobilized peptide,
contacting the receptor with a zslg33 peptide (comprising residues 24-37
of AAB62649), where the receptor binds to the zslg33 peptide. The method
is useful for purifying cells, purifying a peptide, stimulating signal

transduction in a cell expressing a receptor. It is also useful for
modulating secretion of hormones, neural development and/or utilization,
gastric contractility, nutrient uptake, secretion of digestive and
pancreatic enzymes and hormones, secretion of insulin-like growth factor
-I, secretion of non-zslg33 proteins. It is useful for modulating growth
hormone secretion in a mammal having a disease associated with abnormal
levels of growth hormone, such as osteoporosis, bone repair, bone
remodeling, low osteoblast levels, cartilage repair and remodeling,
skeletal dysplasia, immune suppression, obesity, growth retardation,
protein catabolic responses after surgery, cachexia, protein loss,
dwarfism, wound healing and ovulation induction, treating a mammal having
a metabolic disorder requiring neurological feedback, such as satiety
regulation, glucose absorption and metabolism and neuropathy-associated
gastrointestinal disorders, and stimulating glucose-induced insulin
release in a mammal. The present sequence represents the DNA encoding
the long form of motilin receptor, GPR-38A (one of the two isoforms of
GPR38 which result from alternative splicing). GPR38 has homology to the
human G-protein coupled receptor, GHS-R.

Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;

Query Match 77.6%; Score 1078; DB 22; Length 1239;
Best Local Similarity 89.1%; Pred. No. 2e-186;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

```

Qy 1 ATGGGAGAGCCCTGGAAACGACAGCGCCCGAGAGGGGCGGAGCGCGCTGGCCC 60
Db 1 ATGGGAGAGCCCTGGAAACGACAGCGCCCGAGAGGGGCGGAGCGCGCTGGCCC 60
Qy 61 GCGCTGCGGCTTGGACAGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGCGTGGTCCG 120
Db 61 GCGCTGCGGCTTGGACAGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGCGTGGTCCG 120
Qy 121 GTACACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTACACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGAGATCGGGACACCAACCACTTGTACCTGGGCGAG 240
Db 181 ATGCTGATCGGGCGCTACCGGAGATCGGGACACCAACCACTTGTACCTGGGCGAG 240
Qy 241 GCGGTGTCCGACCTACTATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 300
Db 241 GCGGTGTCCGACCTACTATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 300
Qy 301 TCGGCGCCCTGGGTGTTGCGGGCGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGCGCCCTGGGTGTTGCGGGCGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCACTTACGCCACCGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Db 361 TGCACTTACGCCACCGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Qy 421 TCGGCGCCCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCGGCGCCCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GTGCTCTGGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTCTGGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCCTCCG 600
Db 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCCTCCG 600
Qy 601 CCTCTGCGCTGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CCTCTGCGCTGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GGGCCCGAGACCCCGGAGCGCGGCGCTGTTTACGCGCGCAATGCGGCGCGAGCCCGCG 720
Db 661 GGGCCCGAGACCCCGGAGCGCGGCGCTGTTTACGCGCGCAATGCGGCGCGAGCCCGCG 720

```

QY 721 CAGTGGGCGCGTGTGTCATGCTGTGGGTACACCGCCCTACTTCTTCTGCGCCCTTT 780
 Db 721 CAGTGGGCGCGTGTGTCATGCTGTGGGTACACCGCCCTACTTCTTCTGCGCCCTTT 780
 QY 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGGAGCTGTGGAGCAGCCGGCGCG 840
 Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGGAGCTGTGGAGCAGCCGGCGCG 840
 QY 841 CTGGAGGCGCGCGCTCGGGGGAGAGAGGCCACCGGAGACCGTCCGCTGCTG 900
 Db 841 CTGGAGGCGCGCGCTCGGGGGAGAGAGGCCACCGGAGACCGTCCGCTGCTG 900
 QY 901 CGTAAGTGGAGCGCGCTGTGTTCCAAAGACGCTGCTGCAAGTCCGCGCGCGGAGCC 960
 Db 901 C----- 901
 QY 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCCAGCTCTGGGGCGCGCTTCCAGTCCC 1020
 Db 902 ----- 901
 QY 1021 TTTCTATTTCGATTCCAGCTCCACCGCGGTGGTGTCTGGCATTTATATTG 1080
 Db 902 -----TGGTGGTGGTCTTGGCATTTATAATTG 929
 QY 1081 CTGTTGCCCTTCACGTTGGCAGAAATCAITTAATAAACAAGAAATTCGGCGATGAT 1140
 Db 930 CTGTTGCCCTTCACGTTGGCAGAAATCAITTAATAAACAAGAAATTCGGCGATGAT 989
 QY 1141 GTACTTCTCAGTACTTTAAACATGCTGCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
 Db 990 GTACTTCTCAGTACTTTAAACATGCTGCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
 QY 1201 CAACCCAACTCTACAACTCTATTTCAAGAGATACAGAGCGCGGCTTTAACTGCT 1260
 Db 1050 CAACCCAACTCTACAACTCTATTTCAAGAGATACAGAGCGCGGCTTTAACTGCT 1109
 QY 1261 GCTCGAAGGAAGTCCAGAGCGGAGAGGCTTCCACAGAGAGAGGACACTCGCGGGGAAGT 1320
 Db 1110 GCTCGAAGGAAGTCCAGAGCGGAGAGGCTTCCACAGAGAGAGGACACTCGCGGGGAGT 1169
 QY 1321 TGACGGGGAACATGGAGGAGACAGGTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1380
 Db 1170 TGACGGGGAACATGGAGGAGACAGGTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1229
 QY 1381 GATGGGATAA 1390
 Db 1230 GATGGGATAA 1239

RESULT 5

AAF85449 ID AAF85449 standard; cDNA; 1239 BP.

XX AC AAF85449;

XX 23-JUL-2001 (first entry)

XX Nucleotide sequence of a human motilin receptor polypeptide.

XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1239

FT /tag= a

FT /note= "motilin receptor"

XX W0200132710-Al..

XX 10-MAY-2001.

XX PD

PF 25-OCT-2000; 2000WO-US29426.
 XX 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 XX PI Tan C, McKee K;
 XX WPI; 2001-343479/36.
 DR P-PSDB; AAB68478.
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX Disclosure; Page 34; 42pp; English.
 XX The present sequence encodes a human motilin receptor polypeptide.
 CC The specification describes a unique sequence present in exon 1 of
 CC the dog motilin receptor, which is not present in human or Spaeorides
 CC nephelus 7587 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;
 Query Match 77.6%; Score 1078; DB 22; Length 1239;
 Best Local Similarity 89.1%; Pred. No. 2e-186; 0; Indels 151; Gaps 1;
 Matches 1239; Conservative 0; Mismatches 0;
 QY 1 ATGGGAGCGCCCTTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGGTGSCCC 60
 Db 1 ATGGGAGCGCCCTTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGGTGSCCC 60
 QY 61 GCCTGCGCGCTTGCAGAGCGCGCTGCTGCGCCCTTCCCTTGGGGGGCTGTGTCGG 120
 Db 61 GCCTGCGCGCTTGCAGAGCGCGCTGCTGCGCCCTTCCCTTGGGGGGCTGTGTCGG 120
 QY 121 GTGACCGCTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 GTGACCGCTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTTGGGAGCATG 240
 Db 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTTGGGAGCATG 240
 QY 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
 Db 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
 QY 301 TCGCGGCGCTGGGTGTTGCGGCGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTG 360
 Db 301 TCGCGGCGCTGGGTGTTGCGGCGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTG 360
 QY 361 TGACCTTACGCCACCGCTGTGACATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTC 420
 Db 361 TGACCTTACGCCACCGCTGTGACATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTC 420
 QY 421 TCGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 421 TCGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 GTGCTCTGGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 GTGCTCTGGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 CAGGACCCCGGCGCATCTCCGCTAGTCCCGGGCTCAATGGCACCGCGGATCGGCTCTCC 600

Db 541 CAGACCCCGCATCTCGTAGTCCCGGCTCAATGGACCGCGGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCGTCCGCGCTCTCTGCTCTCGCGGCGCACCGCGTCCCGCGCTCG 660
Db 601 CCTCTCGCTCGTCCGCGCTCTCTGCTCTCGCGGCGCACCGCGTCCCGCGCTCG 660
Qy 661 GGGCCGAGACCGGAGGCGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 720
Db 661 GGGCCGAGACCGGAGGCGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 720
Qy 721 CAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGCATCTCTCTAGGCTCATCGGCGGAGCTGTGAGCAGCGCGGCGG 840
Db 781 CTGTGCTCAGCATCTCTCTAGGCTCATCGGCGGAGCTGTGAGCAGCGCGGCGG 840
Qy 841 CTGGAGCGCGCGCTCTGGGCGGAGAGGCGCACCGGCGAGCTGCTGCTGCTGCT 900
Db 841 CTGGAGCGCGCGCTCTGGGCGGAGAGGCGCACCGGCGAGCTGCTGCTGCTGCT 900
Qy 901 CGTAAGTGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 C----- 901
Qy 961 GCGAAACGTGGTCT 1020
Db 902 ----- 901
Qy 1021 TTTCTATTTCGATTCCAGCTCCACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
Db 902 -----TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 929
Qy 1081 CTGTTGCCCTTCCAGCTGCGAGATCATTTACATTAACACGGAAGATTTCGGGATGAT 1140
Db 930 CTGTTGCCCTTCCAGCTGCGAGATCATTTACATTAACACGGAAGATTTCGGGATGAT 989
Qy 1141 GTACTTCTCAGTACTTTAAACATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1200
Db 990 GTACTTCTCAGTACTTTAAACATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1049
Qy 1201 CAACCAATCTCTAACCTCTATTCAAGAGTACAGAGCGCGGCTTTAACTGCT 1260
Db 1050 CAACCAATCTCTAACCTCTATTCAAGAGTACAGAGCGCGGCTTTAACTGCT 1109
Qy 1261 GCTGCAAGAGTCCAGGCGGAGGCTTCCAGAGCAGGACACTCGGCGGAGT 1320
Db 1110 GCTGCAAGAGTCCAGGCGGAGGCTTCCAGAGCAGGACACTCGGCGGAGT 1169
Qy 1321 TGCAGGGGACACTGGAGGAGACGCTGGGCTACACCGAGACAGCGCTAACGTGAAGAC 1380
Db 1170 TGCAGGGGACACTGGAGGAGACGCTGGGCTACACCGAGACAGCGCTAACGTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

RESULT 6
ABK90132
ID ABK90132 standard; DNA; 1239 BP.
XX
AC ABK90132;
XX
DT 21-OCT-2002 (first entry)
XX
DE DNA encoding human G protein-coupled receptor 38 (GPR38).
XX
KW Human; G protein-coupled receptor 38; receptor; GPR38; gene; da;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;

KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX Homo sapiens.
PH Location/Qualifiers
FT 1..1239
FT /tag= a
FT /product= "Human G protein-coupled receptor 38 (GPR38)"
PN WO200257791-A2.
XX
XX 25-JUL-2002.
XX 29-NOV-2001; 2001WO-US45219.
XX 29-NOV-2000; 2000US-250251P.
PR 30-NOV-2000; 2000US-250452P.
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX Brown JP, Burmer GC, Roush CL, Kulander BG;
XX WPI; 2002-566812/60.
XX P-PSDB; ABG30936.
XX Assay for detecting Alzheimer's disease, Parkinson's disease,
XX ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
XX carcinoma, comprises using a binding partner for G protein coupled
XX receptor 38 -
XX Disclosure; Fig 1; 112pp; English.
XX The present invention relates to a new assay method that involves
XX contacting a binding partner specific for G protein coupled receptor
XX (GPR) 38 with specific cells. The method of the invention is useful for
XX the detection of an increased risk of Alzheimer's disease, Parkinson's
XX disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
XX glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
XX for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
XX disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
XX glioblastoma, breast carcinoma, colon carcinoma, lung small cell
XX carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
XX pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
XX to manufacture a medicament able to reduce the symptoms of these
XX diseases. Nucleic acids encoding GPR 38 can also be used to treat the
XX diseases. The present nucleic acid sequence encodes the human G protein-
XX coupled receptor 38 (GPR38) of the invention.
XX SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;

Query Match 77.6%; Score 1078; DB 24; Length 1239;
Best Local Similarity 89.1%; Pred. No. 2e-186;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
Qy 1 ATGGCGAGCCCTTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGTGGCCC 60
Db 1 ATGGCGAGCCCTTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGTGGCCC 60
Qy 61 GCGTCCGCGCTTGGACAGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 GCGTCCGCGCTTGGACAGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Qy 121 GTACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGACATCGCGACACCAACCACTTGTACCTGGGAGCATG 240
Db 181 ATGCTGATCGGGCGCTACCGGACATCGCGACACCAACCACTTGTACCTGGGAGCATG 240
Qy 241 GCGGTGTCGAGCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GCGGTGTCGAGCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 301 TCGCGGCGCTGGTGTTCGGGCGGCTGTCTGCGCGCTGTCCCTCTACGTGGGCGAGGCG 360
 Db 301 TCGCGGCGCTGGTGTTCGGGCGGCTGTCTGCGCGCTGTCCCTCTACGTGGGCGAGGCG 360
 QY 361 TGCACCTACGACCGCTGTGACATGACCGCGCTCAGCGTGCAGCGCTACCTGGCCATC 420
 Db 361 TGCACCTACGACCGCTGTGACATGACCGCGCTCAGCGTGCAGCGCTACCTGGCCATC 420
 QY 421 TCGCGCGGCTCGCGCGGCTGTGTCACCGCGCGCGCTCGCGCGCTCATCGCT 480
 Db 421 TCGCGCGGCTCGCGCGGCTGTGTCACCGCGCGCGCTCGCGCGCTCATCGCT 480
 QY 481 GTGCTCTGGGCGGTGGGCGCTGTCTCTGCGCGCTCTCTGTCGTGGGCGGTGCGAG 540
 Db 481 GTGCTCTGGGCGGTGGGCGCTGTCTCTGCGCGCTCTCTGTCGTGGGCGGTGCGAG 540
 QY 541 CAGACCCCGGCACTCTCGTAGTCCGGGCTCAATGGCAACCGCGGATCGCTCTCTCG 600
 Db 541 CAGACCCCGGCACTCTCGTAGTCCGGGCTCAATGGCAACCGCGGATCGCTCTCTCG 600
 QY 601 CCTCTCGGCTGTGCGCGCTCTCTGCTCTCGGCGGCGCACCGCGCTCCGCGCTCG 660
 Db 601 CCTCTCGGCTGTGCGCGCTCTCTGCTCTCGGCGGCGCACCGCGCTCCGCGCTCG 660
 QY 661 GGGCGCGAGACCGCGAGGCGCGCGCTGTCTCAGCGCGCAATGCGCGCGAGCGCGCG 720
 Db 661 GGGCGCGAGACCGCGAGGCGCGCGCTGTCTCAGCGCGCAATGCGCGCGAGCGCGCG 720
 QY 721 CAGTGGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 780
 Db 721 CAGTGGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 780
 QY 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGCGGAGCTGTGGAGCAGCGCGCGCG 840
 Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGCGGAGCTGTGGAGCAGCGCGCGCGCG 840
 QY 841 CTGCGAGCGCGCGCGCTCGGCGGCGAGAGGCGCACCGCGAGCGCTCGGCTCGTG 900
 Db 841 CTGCGAGCGCGCGCGCTCGGCGGCGAGAGGCGCACCGCGAGCGCTCGGCTCGTG 900
 QY 901 CGTAAGTGGAGCGCGCTGTGTTCAAGACGCGCTGTGCGAGTCCGCGCGCGCGGAGC 960
 Db 901 C----- 901
 QY 961 GCGCAACGCTGGTGTCCCTCTCTGCGCGAGCTGTGGCGCGCTTCCAGCTCCC 1020
 Db 902 ----- 901
 QY 1021 TTTCCTATTTCGATTCAGCGCTCCACCGCGGTGGTGTGCTGTGCAATTTAATTG 1080
 Db 902 ----- 901
 QY 1081 CTGTTGCGCTTCCAGTGTGGCAGAACTCAATTAATAAACAAGGATTCGCGGATGAT 1140
 Db 930 CTGTTGCGCTTCCAGTGTGGCAGAACTCAATTAATAAACAAGGATTCGCGGATGAT 999
 QY 1141 GTACTTCTCAGTACTTTAATCATGCTGTGCTGTGCACTTTTCTATCTGAGCGCATCTAT 1200
 Db 990 GTACTTCTCAGTACTTTAATCATGCTGTGCTGTGCACTTTTCTATCTGAGCGCATCTAT 1049
 QY 1201 CAACCCCAATCTCTACAACTCTATTTCAAAGAGTACAGAGCGCGCGCTTTAACTGCT 1260
 Db 1050 CAACCCCAATCTCTACAACTCTATTTCAAAGAGTACAGAGCGCGCGCTTTAACTGCT 1109
 QY 1261 GCTCGCAAGAGTCTCAGGCGGAGGCTTCCACAGAGCAGGAGACATCGCGGGGAGT 1320
 Db 1110 GCTCGCAAGAGTCTCAGGCGGAGGCTTCCACAGAGCAGGAGACATCGCGGGGAGT 1169
 QY 1321 TGCAGGGGACACTGGAGGAGACACGTTGGGCTACCGGACAGCGCTAACGTGAAGAC 1380
 Db 1170 TGCAGGGGACACTGGAGGAGACACGTTGGGCTACCGGAGCAGCGCTAACGTGAAGAC 1229

QY 1381 GATGGGATAA 1390
 Db 1230 GATGGGATAA 1239
 RESULT 7
 ABZ42842
 ID ABZ42842 standard; DNA; 1239 BP.
 XX AC ABZ42842;
 XX DT 04-MAR-2003 (first entry)
 XX Human motilin receptor GPR38 nucleotide SBQ ID NO:472.
 XX
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer; gene; ds.
 XX Homo sapiens.
 XX WO200261087-A2.
 XX 08-AUG-2002.
 XX 19-DEC-2001; 2001WO-US50107.
 XX 19-DEC-2000; 2000US-257144P.
 XX (LIFE-) LIPESPAN BIOSCIENCES INC.
 XX Burner GC, Roush CL, Brown JP;
 WPI; 2003-046718/04.
 P-PSDB; ABP81993.
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 (GPCR), useful for diagnosing and designing drugs for treating
 conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
 cancer or autoimmune diseases -
 XX Disclosure; Fig 1; 523pp; English.
 The present invention describes antigenic peptides (I) comprising:
 (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 acids. Also described: (1) an assay for the detection of a particular
 G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 and (2) an isolated antibody having high specificity and high affinity
 or avidity for a particular GPCR. (I) can be used as GPCR modulators and
 in gene therapy. The antigenic peptides for GPCRs are useful in detecting
 an antibody against a particular GPCR, and in the production of specific
 antibodies. The peptides and antibodies are also useful for detecting the
 presence or absence of corresponding GPCRs. The antigenic peptides for
 GPCRs and antibodies are useful for diagnosing and designing drugs for
 treating immune-related diseases, growth-related diseases, cell
 regeneration-related disease, immunological-related cell proliferative
 diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 atherosclerosis, bacterial, fungal, protozoan or viral infections,
 osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 inflammation, allergies, Crohn's disease, diabetes, graft versus host
 disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 anxiety, depression, schizophrenia, dementia, mental retardation, memory
 loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 any other disorder in which GPCRs are involved. The antibodies may be


```

PR 20-NOV-1998; 98US-0109213.
PR 27-NOV-1998; 98US-0110060.
PR 16-FEB-1999; 99US-0120416.
PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123944.
PR 12-MAR-1999; 99US-0123945.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123948.
PR 12-MAR-1999; 99US-0123949.
PR 12-MAR-1999; 99US-0123951.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 28-MAY-1999; 99US-0137567.
PR 30-JUN-1999; 99US-0141448.
PR 27-AUG-1999; 99US-0151114.
PR 03-SEP-1999; 99US-0152524.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156634.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Behan DP, Lehmann-Bruinema K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX
DR WPI: 2000-317986/27.
DR P-PSDB; AAB02854.
XX
PT Non-endogenous, human G protein-coupled receptors for screening
PT receptor, inverse or partial agonists useful as therapeutic agents -
XX
PS Example 2; Page 166-168; 187pp; English.
XX
CC The present invention describes transmembrane receptors, preferably
CC human G protein coupled receptors (GPCR), for which the endogenous
CC ligand is unknown (orphan GPCR receptors). More specifically the present
CC invention relates to non-endogenous, constitutively activated versions
CC of a human GPCR. These non-endogenous human GPCRs can be useful for
CC the direct identification of candidate compounds as receptors agonists,
CC inverse agonists or partial agonists for use as pharmaceutical agents.
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC the exemplification of the present invention.
XX
SQ Sequence 2040 BP; 350 A; 630 C; 597 G; 463 T; 0 other;

Query Match 75.3%; Score 1047.2; DB 21; Length 2040;
Best Local Similarity 99.7%; Pred. No. 7.9e-181;
Matches 1049; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGAGCCCTGGAAACGCGAGCGGCGCCCGGAGGGGCGGGAGCGCGTGGCCC 60
DB 1 ATGGGAGCCCTGGAAACGCGAGCGGCGCCCGGAGGGGCGGGAGCGCGTGGCCC 60
QY 61 GCGCTGCGCGCTTGGACGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGCGCTGGTCCG 120
DB 61 GCGCTGCGCGCTTGGACGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGCGCTGGTCCG 120
QY 121 GTACCGCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 GTACCGCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 ATGCTGATCGGGGCTACCGGGACATCGGGACACCAACCACTTGTACCTTGGGAGCATG 240
DB 181 ATGCTGATCGGGGCTACCGGGACATCGGGACACCAACCACTTGTACCTTGGGAGCATG 240
QY 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTTCCGACCTGTACCGCTTGGCGC 300
DB 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTTCCGACCTGTACCGCTTGGCGC 300
QY 301 TCGGGGCGCTGGGTGTTTCGGGGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360

```

RESULT 10

AAZ45402

ID AAZ45402 standard; DNA; 3066 BP.

XX

AC AAZ45402;

XX

DT 27-MAR-2000 (first entry)

XX

Genomic sequence of the motilin receptor gene including 5' UTR.

DE

XX

XX

KW

KW

KW

KW

KW

KW

KW

KW

KW

XX

Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 spliced form; MTL-R1a; MTL-R1b; gastric motility disorder;
 functional defect; neurological disorder; scleroderma; colonoscopy;
 paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 infection; stress-related motility disorder; psychogenic disorder;
 gastroparesis; gastro-oesophageal reflux disease; constipation;
 chronic idiopathic pseudo obstruction; acute faecal impaction;
 postoperative ileus; gallstones; infantile colic; diarrhoea;
 irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 endoscopy; duodenal intubation; db.

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 1929..1931
 FT misc_feature /*tag= a
 FT /*note= "imperfect donor site"
 FT intron 1930..2728
 FT /*tag= b
 FT /*note= "intronic sequence"
 FT 2080..2082
 FT /*tag= c
 FT /*note= "perfect donor site"
 FT 2729..2732
 FT /*tag= d
 FT /*note= "perfect splice acceptor site"
 XX
 PN W09964436-A1.
 XX
 XX 16-DEC-1999.
 XX
 XX 08-JUN-1999; 99WO-US12773.
 XX
 XX 12-JUN-1998; 98US-0089098.
 XX
 XX (MERI) MERCK & CO INC.
 XX
 XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;
 XX
 XX WPI; 2000-105868/09.
 XX
 XX Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions
 XX
 PS Example 1; Fig 1; 44pp; English.
 XX
 CC The present sequence represents the genomic sequence of the motilin
 CC receptor gene, including the 5' untranslated region (5' UTR). This gene
 CC encodes a G-protein coupled receptor, and is designated MTL-R1 (also
 CC GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see AAY54145) and
 CC MTL-R1B (see AAY54146). MTL-R1A is a functional seven transmembrane
 CC domain form, and MTL-R1B is a truncated five transmembrane domain. The
 CC MTL-R1 proteins are used to identify agonists and antagonists which can
 CC be used for treating gastric motility disorders, functional defects,
 CC disorders secondary to neurological disorders e.g. scleroderma,
 CC paraneoplastic syndromes radiation induced dysmotility, diabetes,
 CC infections, stress-related motility disorders, psychogenic disorders,
 CC gastroparesis, gastro-oesophageal reflux disease, constipation,
 CC chronic idiopathic pseudo obstruction, acute faecal impaction,
 CC postoperative ileus, gallstones, infantile colic, irritable bowel
 CC syndrome, non-ulcer dyspepsion, non-cardiac chest pain and diarrhoea.
 CC They can also be used in the preparation for colonoscopy, endoscopy and
 CC duodenal intubation.
 XX
 SQ Sequence 3066 BP; 585 A; 930 C; 926 G; 625 T; 0 other;
 Query Match 74.7%; Score 1038; DB 21; Length 3066;
 Best Local Similarity 99.8%; Pred. No. 3.8e-179;
 Matches 1050; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 1 ATGGGAGACCCCTGGAAACGACGACGCGCCGCGGGGCGCGGAGCGCGCTGGCCC 60
 DB 1029 ATGGGAGACCCCTGGAAACGACGACGCGCCGCGGGGCGCGGAGCGCGCTGGCCC 1088
 QY 61 GCCTGCGCGCTTGGACGACGCGCGCTGCTCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
 DB 1089 GCCTGCGCGCTTGGACGACGCGCGCTGCTCGCCCTTTCCCTGGGGGCGCTGGTGGCG 1148
 QY 121 GTACACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 1149 GTACACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1208
 QY 181 ATGCTGATCGGGCGCTACCGGGGACATGCGGGACACACCACTTGTACCTGGGCGAGCATG 240

DB 1209 ATGCTGATCGGGCGCTACCGGGACATGCGGGACACCAACCACTTGTACCTGGGCGAGCATG 1268
 QY 241 GCGGTGTCGACCTACTCTCTGCTCGGGCTGCGCTTGCACCTGTACCGCTTGTGGCGC 300
 DB 1269 GCGGTGTCGACCTACTCTCTGCTCGGGCTGCGCTTGCACCTGTACCGCTTGTGGCGC 1328
 QY 301 TCGCGGCTTGGGTGTTGCGGGCGCTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 1329 TCGGGGCTTGGGTGTTGCGGGCGCTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1388
 QY 361 TGCACCTACGCGACCTGTGACATGACCGGCTCAGCGTTCAGCGCTACCTGGGCGATC 420
 DB 1389 TGCACCTACGCGACCTGTGACATGACCGGCTCAGCGTTCAGCGCTACCTGGGCGATC 1448
 QY 421 TGCAGCGCTTGGGTGTTGCGGGCGCTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 1449 TGCAGCGCTTGGGTGTTGCGGGCGCTGCTTGGTCACTGCGGCGCTGCTGCGGCTCATGCT 1508
 QY 481 GTGCTCTGGGCGCTGCGGCTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 DB 1509 GTGCTCTGGGCGCTGCGGCTGCTTCTGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1568
 QY 541 CAGGACCCCGGCTTCCGTAGTCCCGGCTTCAATGGCACCCGCGGATTCGCTCTCTCG 600
 DB 1569 CAGGACCCCGGCTTCCGTAGTCCCGGCTTCAATGGCACCCGCGGATTCGCTCTCTCG 1628
 QY 601 CTTCTGCTGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 1629 CTTCTGCTGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1688
 QY 661 GGGCCGAGACCGCGGAGCGCGGCTGTTTCAGCGCGGAAATGCGCGCGGAGCCCGCG 720
 DB 1689 GGGCCGAGACCGCGGAGCGCGGCTGTTTCAGCGCGGAAATGCGCGCGGAGCCCGCG 1748
 QY 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 1749 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1808
 QY 781 CTGTGCTCAGCATCTCTACCGGCTCATCGGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 840
 DB 1809 CTGTGCTCAGCATCTCTACCGGCTCATCGGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 1868
 QY 841 CTGCGAGGCGCGCGCTGCGGCGGAGAGGCGCACCGGAGAGCTGCGGAGCTGCGGAGCTGCTG 900
 DB 1869 CTGCGAGGCGCGCGCTGCGGCGGAGAGGCGCACCGGAGAGCTGCGGAGAGCTGCGGAGCTGCTG 1928
 QY 901 CGTAAGTGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 1929 CGTAAGTGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1988
 QY 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 1989 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2048
 QY 1021 TTTCTATTTCGATTCCAGCTCCACCCCGCG 1052
 DB 2049 --TCCTATTTCGATTCCAGCTCCACCCCGCG 2078
 RESULT 11
 ID AAF85448 standard; cDNA; 1203 BP.
 XX
 AC AAF85448;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Nucleotide sequence of a rabbit motilin receptor polypeptide.
 XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea; ss.
 XX

XX AAF85447;
AC 23-JUL-2001 (first entry)
DT Nucleotide sequence of a dog motilin receptor exon 1.
DE Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea; ss.
KW Canis sp.
OS
FH Key Location/Qualifiers
FT 1..813
FT /*tag= a
FT /note= "motilin receptor exon 1"
XX WO200132710-A1.
XX 10-MAY-2001.
XX 25-OCT-2000; 2000WO-US29426.
XX 29-OCT-1999; 99US-0162264.
XX (MERI) MERCK & CO INC.
XX Tan C, McKee K;
PI WPI; 2001-343479/36.
DR P-PSDB; AAB68476.
XX Novel polypeptides related to dog and rabbit motilin receptor
PT polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX Claim 14; Page 17-18; 42pp; English.
XX The present sequence represents exon 1 of a dog motilin receptor gene.
CC The specification describes a unique sequence present in exon 1 of
CC the motilin receptor, which is not present in human or Sphaeroides
CC naphelus 7587 motilin receptor sequences. The unique nucleic acid
CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.
XX SQ Sequence 813 BP; 65 A; 316 C; 305 G; 127 T; 0 other;
Query Match 39.3%; Score 546; DB 22; Length 813;
Best Local Similarity 80.5%; Pred. No. 4.5e-90;
Matches 705; Conservative 0; Mismatches 105; Indels 66; Gaps 3;
QY 25 GAGCGCCCGAGGGGGCGGGAGCCCGCTGGCCCGCTGCGCGTTCGACGAGCGC 84
DB 4 GGGGGCCCGGGAACAGCAGCGAGCGGGCGCGAGCTGCCGTGCGACGAGCGC 63
QY 85 CGCTGCTCGCCCTTCCCTCGGGGGCGCTGTGCGGGTACCGCTGTGCTGCTG 144
DB 64 CTGTGCTGCGCCCTTCCCTCGGGGGCGCTGTGCGGGTACCGCTGTGCTGCTG 123
QY 145 TTGCTGCTCGGGGTGAGCGCAACGCTGTGACCGTGTGATGCTGGGGCGCTACCGGGAC 204
DB 124 TTGCGCGTGGCGGTGAGCGCAACCTGTGTGACCGTGTGCTGTGACCGCGTACCGGGAC 183
QY 205 ATGGGACACACCAACCTGTGTACCTGGGAGCAGTGGCGGTGCTGCTGCTGCTG 264
DB 184 ATGGCACCACCAACCTGTGTACCTGGGAGCAGTGGCGGTGCTGCTGCTGCTG 243
QY 265 CTGGGGTGGCTTTCGACCTGTACCGCTCTTGGGGCTCGGGCGCTTGGGGCGG 324

DB 244 CTGGGGCTGCCCTCGACCTGTACCGCTGTGGCGCTCGCGCCCTGGGTGTTTCGGGCAG 303
QY 325 CTGCTCTGCGCGCTGTCCCTCTACGTGGGGAGGGTGCACCTACGCCACGCTGCTGCAC 384
DB 304 CTGCTGTGCGCGCTGTGCTGTACCTGGGGAGGGTGCACCTACGCCACGCTGCTGCAC 363
QY 385 ATGACCGCTCAGCGCTCGAGCGCTACCTGGCCATCTGCGCGCCGCTCGCGCCCGCGTC 444
DB 364 GTGACCGCTGAGCGCTGAGCGCTACCTGGCCGCTGTGCGCGCCGCTCGCGCCCGCGG 423
QY 445 TTGGTACCGCGCGCGCGCTTCGCGCGCTCATCGCTGTGCTCTGGGGCGTGGGCTGCTC 504
DB 424 CTGCTGTGCGCGCGCGCGCGCTCATCGCGCGCTCTGGGGCGTGGGCTGCTGCTG 483
QY 505 TCTGCCGCTCCCTTCTTGTTCCTGGTGGGGTGCAGCAGGACCCCGGCATCTCGGTAGTC 564
DB 484 TCGGCGCGCGCTTCTTCTTCTGGTGGGGTGCAGCAGGACCGGCGC-----GGC 534
QY 565 CCGGGCTCAATGGCACCGCGGATCGCCTCTCGCCTCTCGCCTCGTCCGCGCTCTC 624
DB 535 CCGGGCTCAACGGCAGCGCGC----- 557
QY 625 TGCTCTCGCGCGCGCACCGCGCTCCCGCGTTCGGGGCCCGAGACCGCGAGGCGCG 684
DB 558 -----GCTGGCGCGCGCGCTCCCGCGCGCGCGCGCG-----GCG 597
QY 685 GCGCTGTTCAGCGCGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGTGCTG 744
DB 598 GCGCTCTTCAGCGCGGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGT 657
QY 745 CTGCGGTACACCGCGCTACTTCTTCGCGCTTCTGCTGCTCAGCATCTCTACGCG 804
DB 658 CTGCGGTACACCGCGCTACTTCTTCGCGCTTCTGCTGCTCAGCATCTCTGCTG 717
QY 805 CTGCGGTACACCGCGCTACTTCTTCGCGCTTCTGCTGCTCAGCATCTCTGCTG 864
DB 718 CGCATCGCGCGCGAGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGG 777
QY 865 CGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 900
DB 778 CGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 813
RESULT 13
ABQ47146/c
ID ABQ47146 standard; DNA; 1179 BP.
XX AC ABQ47146;
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 33737.
XX DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP10074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 03-SEP-2000; 2000DE-104543.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:43:42 ; Search time 98.3416 Seconds
(without alignments)
6238.690 Million cell updates/sec

Title: US-09-719-485-4

Perfect score: 1390

Sequence: 1 atgggcagcccttggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	283	20.4	283	3	US-08-993-088A-4
2	283	20.4	283	4	US-08-993-424B-4
3	283	20.4	283	4	US-09-603-680-4
4	234.8	16.9	1063	3	US-09-077-675A-1
5	234.8	16.9	1063	4	US-09-077-674-1
6	230	16.5	1029	3	US-09-077-675A-4
7	230	16.5	1029	4	US-09-077-674-4
8	229	16.5	250	4	US-09-016-434-359
9	228.8	16.5	1122	3	US-09-077-675A-9
10	228.8	16.5	1122	4	US-09-077-674-9
11	228.4	16.4	1092	3	US-09-077-675A-15
12	228.4	16.4	1092	4	US-09-077-674-15
13	228.4	16.4	3129	3	US-09-077-675A-14
14	228.4	16.4	3129	4	US-09-077-674-14
15	227.8	16.4	1088	3	US-09-077-675A-6
16	227.8	16.4	1088	4	US-09-077-674-6
17	227.8	16.4	1101	4	US-09-016-434-1148
18	227.8	16.4	1101	4	US-09-170-496D-87
19	227.8	16.4	1101	4	US-09-170-496D-209
20	153	11.0	836	3	US-09-077-675A-11
21	153	11.0	836	4	US-09-077-674-11
22	134.8	9.7	1248	4	US-09-545-944-1
23	127	9.1	1575	3	US-08-858-876A-1
24	127	9.1	1575	3	US-09-472-880-1
25	125.4	9.0	1342	3	US-08-832-399-1
26	125.4	9.0	1342	3	US-09-372-498-1
27	123	8.8	4080	4	US-09-016-434-1346

Sequence 3, Appl1
Sequence 3, Appl1
Sequence 1366, Ap
Sequence 12, Appl
Sequence 113, App
Sequence 223, App
Sequence 3, Appl1
Sequence 1, Appl1
Sequence 6, Appl1
Sequence 6, Appl1
Sequence 27, Appl1
Sequence 3, Appl1
Sequence 18, Appl
Sequence 19, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-08-993-088A-4
; Sequence 4, Application US/08993088A
; Patent No. 6287855
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,088A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1....283
; OTHER INFORMATION: cdna probe
; US-08-993-088A-4

28 122.2 8.8 1529 3 US-08-858-876A-3
29 122.2 8.8 1529 3 US-09-472-880-3
30 122 8.8 1285 4 US-09-016-434-1366
31 120 8.6 1535 4 US-09-668-680-12
32 119 8.6 1212 4 US-09-170-496D-113
33 119 8.6 1212 4 US-09-170-496D-223
34 115.4 8.3 1276 4 US-09-200-090-3
35 115.4 8.3 1233 4 US-09-200-090-1
36 109.8 7.9 1164 3 US-08-993-088A-6
37 109.8 7.9 1164 4 US-08-993-424B-6
38 109.8 7.9 1164 4 US-09-603-680-6
39 109.8 7.9 1365 4 US-08-899-112B-27
40 108.2 7.8 1219 4 US-08-981-700A-3
41 106.8 7.7 1116 3 US-08-993-088A-18
42 106.8 7.7 1116 3 US-08-993-088A-19
43 106.8 7.7 1116 4 US-08-993-424B-18
44 106.8 7.7 1116 4 US-09-603-680-18
45 106.8 7.7 1116 4 US-09-603-680-19

Query Match 20.4%; Score 283; DB 3; Length 283;
 Best Local Similarity 100.0%; Pred. No. 2.6e-48;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCAACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCCTGC 265
 Db 1 TGGGACCAACCAACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCCTGC 60

QY 266 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGCTCGCGCCCTGGGTCGCGGCCG 325
 Db 61 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGCTCGCGCCCTGGGTCGCGGCCG 120

QY 326 TGCTCTGCCGCTGTCCCTCTACGTGGGCGAGGCTGACCTAGCCACGCTGTGCACA 385
 Db 121 TGCTCTGCCGCTGTCCCTCTACGTGGGCGAGGCTGACCTAGCCACGCTGTGCACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTACCTGGGCATCTGCCGCGCTCCGCGCCGCGTCT 445
 Db 181 TGACCGGCTCAGGCTCGAGGCTACCTGGGCATCTGCCGCGCTCCGCGCCGCGTCT 240

QY 446 TGGTACCCGCGCGCGTCCGCGCTCATCGCTGTGCTCTG 488
 Db 241 TGGTACCCGCGCGCGTCCGCGCTCATCGCTGTGCTCTG 283

RESULT 2

US-08-993-424B-4
 ; Sequence 4, Application US/08993424B
 ; Patent No. 6337206
 ; GENERAL INFORMATION:
 ; APPLICANT: Tan, Carina
 ; APPLICANT: Kolakowski, Lee F., Jr.
 ; TITLE OF INVENTION: MOUSE GALANIN RECEPTOR GALR2 AND
 ; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/993,424B
 ; FILING DATE: 18-DEC-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/033,851
 ; FILING DATE: 27-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heber, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 19846NP2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1958
 ; TELEFAX: 732-594-4720
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 283 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Other
 ; FEATURE:
 ; NAME/KEY: Other
 ; LOCATION: 1...283
 ; OTHER INFORMATION: cdna probe

US-08-993-424B-4
 Query Match 20.4%; Score 283; DB 4; Length 283;
 Best Local Similarity 100.0%; Pred. No. 2.6e-48;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCAACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCCTGC 265
 Db 1 TGGGACCAACCAACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCCTGC 60

QY 266 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGCTCGCGCCCTGGGTCGCGGCCG 325
 Db 61 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGCTCGCGCCCTGGGTCGCGGCCG 120

QY 326 TGCTCTGCCGCTGTCCCTCTACGTGGGCGAGGCTGACCTAGCCACGCTGTGCACA 385
 Db 121 TGCTCTGCCGCTGTCCCTCTACGTGGGCGAGGCTGACCTAGCCACGCTGTGCACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTACCTGGGCATCTGCCGCGCTCCGCGCCGCGTCT 445
 Db 181 TGACCGGCTCAGGCTCGAGGCTACCTGGGCATCTGCCGCGCTCCGCGCCGCGTCT 240

QY 446 TGGTACCCGCGCGCGTCCGCGCTCATCGCTGTGCTCTG 488
 Db 241 TGGTACCCGCGCGCGTCCGCGCTCATCGCTGTGCTCTG 283

RESULT 3

US-09-603-680-4
 ; Sequence 4, Application US/09603680
 ; Patent No. 6544753
 ; GENERAL INFORMATION:
 ; APPLICANT: Tan, Carina
 ; APPLICANT: Sullivan, Kathleen
 ; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/603,680
 ; FILING DATE: 26-Jun-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/033,851
 ; FILING DATE: 27-DEC-1996
 ; APPLICATION NUMBER: 08/993,088
 ; FILING DATE: 18-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heber, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 19846 CA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1958
 ; TELEFAX: 732-594-4720
 ; TELEX: <Unknown>
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 283 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Other

FEATURE:
NAME/KEY: Other
LOCATION: 1...283
OTHER INFORMATION: cDNA probe
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-603-680-4

Query Match 20.4%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.6e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TCGGACCAACCACCACTTGTACCTGGCAGCATGGCGCTTCGCAGACTTATCATCTCTGC 265
DB 1 TCGGACCAACCACCACTTGTACCTGGCAGCATGGCGCTTCGCAGACTTATCATCTCTGC 60

QY 266 TCGGGTGCCGTTTCGACCTGTACCGCTCTGGCGCTTCGCCGCTTCGGGTGTTTCGGGCGC 325
DB 61 TCGGGTGCCGTTTCGACCTGTACCGCTCTGGCGCTTCGCCGCTTCGGGTGTTTCGGGCGC 120

QY 326 TGCTCTGCCGCTTCCTCTACGTGGGAGGGGTGCACCTAGCCACGCTGTCTGCACA 385
DB 121 TGCTCTGCCGCTTCCTCTACGTGGGAGGGGTGCACCTAGCCACGCTGTCTGCACA 180

QY 386 TCAGCGCGCTCAGCTTCGAGCGCTACCTGGCCATCTCGCGCGCTTCGCCGCGCTCT 445
DB 181 TCAGCGCGCTCAGCTTCGAGCGCTACCTGGCCATCTCGCGCGCTTCGCCGCGCTCT 240

QY 446 TGTGTACCCGCGCGCTTCGCCGCTTCATCGCTGTCTGTG 488
DB 241 TGTGTACCCGCGCGCTTCGCCGCTTCATCGCTGTCTGTG 283

RESULT 4
US-09-077-675A-1
Sequence 1, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1063 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-675A-1

Query Match 16.9%; Score 234.8; DB 3; Length 1063;
Best Local Similarity 69.3%; Pred. No. 1.4e-38;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 92 CGCCCTTTCCCCTGGGGCGCTGGTCCGGTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 151
DB 69 CGCTCTTTCCCACGCGCTGTGGGGCGGTACCGCACCTGTGGCGGTCTTCTGCTGTGCTGTG 128

QY 152 TCGGGGTGAGCGGCAACCTGTGTGATCGGTCGTGATCGGCGCTTACCGGACATGCGGA 211
DB 129 TGGGTATCGCGGGCAACCTGTGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGATCGG 188

QY 212 CCACCAACCACTGTGTGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGAT 271
DB 189 CCACCAACCACTGTGTGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGAT 248

QY 272 TCGCGTTTGACCTGTGTGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGATCGGTCGTG 331
DB 249 TCGCGTTTGACCTGTGTGTGATCGGTCGTGATCGGTCGTGATCGGTCGTGATCGGTCGTG 308

QY 332 GGCGCTGTGCTCTACCTGTGGCGAGGCTGTGACCTAGCCACCTGTGTCACATGACCG 391
DB 309 GCNAACCTCTTCAGTTCGTAGGAGAGCTGTGACCTAGCCACCTGTGTCACATGACCG 368

QY 392 CGCTCAGCGTCGAGCGCTACCTGTGGCCATCTGCGCGCGCTTCCGGCGCGCTTCTTGTGTC 451
DB 369 CGCTCAGCGTCGAGCGCTACCTGTGGCCATCTGCGCGCGCTTCCGGCGCGCTTCTTGTGTC 428

QY 452 CCGCGCGCGCGCTTCCGGCGCGCTTCCGGCGCGCTTCCGGCGCGCTTCCGGCGCGCTTCCGG 511
DB 429 CCAAGGCGCGGTAAGCTGTGTCATCTGGGTGTCATCTGGGTGTCATCTGGGTGTCATCTGG 488

QY 512 GTCCCTTCTTGTTCCTGTGGCGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 553
DB 489 GGCCCCATCTTCGTGTGTCGTGGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGGAG 530

RESULT 5
US-09-077-674-1
Sequence 1, Application US/09077674
Patent No. 6531314
GENERAL INFORMATION:
APPLICANT: Arena, Joseph F.
APPLICANT: Cully, Doris F.
APPLICANT: Feighner, Scott D.
APPLICANT: Howard, Andrew D.
APPLICANT: Liberator, Paul A.
APPLICANT: Schaeffer, James M.
APPLICANT: Van Der Ploeg, Leonardus
TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,674
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

Patent No. 6531314
GENERAL INFORMATION:
APPLICANT: Arena, Joseph P.
APPLICANT: Cully, Doris F.
APPLICANT: Feighner, Scott D.
APPLICANT: Howard, Andrew D.
APPLICANT: Liberator, Paul A.
APPLICANT: Schaeffer, James M.
APPLICANT: Van Der Ploeg, Leonardus
TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: US/09/077,674
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19589P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-077-674-4

Query Match 16.5%; Score 230; DB 4; Length 1029;
Best Local Similarity 68.6%; Pred. No. 1.3e-37;
Matches 317; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 92 CGCCCTTCCCTGGGGGCGCTGGTGGCGGTGACCGTGTGCGTGTGCTGTGCTGTGCTG 151
DB 266 CGCTCTCCCGACGCGCTGTGGCGGCGTCCCGCCACCTCGGTGGCGCTTCTG 325
QY 152 TCGGGGTGACGGCAAGCTGGTGTGACCGTGTGATGTCGGCGGTACCGGACATGGGA 211
DB 326 TGGGTATCGGGGCAACTCTCTACGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 385
QY 212 CCACCAACCACTTGTACTTGGGAGCATGCGCGTGTCCGACCTACTCTCTGCTCGGGC 271
DB 386 CCACCAACCACTTGTACTTGGGAGCATGCGCGTGTGTGTGTGTGTGTGTGTGTGTGT 445
QY 272 TGGCGTTCGACCTGTACCGCTCTGGGCGCTTGGGCGCTTGGGCGCTTGGGCGCTTGT 331
DB 446 TGGCGCTTCGAACTCTTCCGCTTTGGCAGTACCGGCGCTTGGAACTTTGGAACTTGT 505
QY 332 GCGCGCTGTCCCTCTACGTGGGAGGCTGCACCTACGCGACGCTGCTGCACATGACCG 391
DB 506 GCAACTCTTCAGTTCGTTAGGAGAGCTGTGACCTACGCGACAGTGTCTCACCATACCG 565
QY 392 CGCTACGCGTCAAGCGCTACTGTGCGCATCTGCGCGCGCTTCCGCGCGCGCTTGTGTCA 451

DB 566 CGCTGAGCGTCGAGCGCTACTTCGCCCATCTGCTTCCCGCTGCGGGCAAGGTAGTGCTCA 625
QY 452 CCGGCGCGCGCGTCCGCGCGCTCATCGCTGTGCTCTGGGCGCGCTGCTGCTCTCTGCGCG 511
DB 626 CCAAGGGCGGGTAAAGCTGGTCTGCTCATCTGGGCGCGTGGCTTCTGCGAGCGCG 685
QY 512 GTCCCTTCTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 553
DB 686 GGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
RESULT 8
US-09-016-434-359
Sequence 359, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 359:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THE1NOT01
CLONE: 2018536
US-09-016-434-359

Query Match 16.5%; Score 229; DB 4; Length 250;
Best Local Similarity 98.8%; Pred. No. 1.6e-37;
Matches 240; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1148 TCTCAGTACTTTAACTGATGCTGCTCTGCAACTTTTCTATCTGAGCGCATCTATCAACCA 1207
DB 1 TCTCAGTACTTTAACTGATGCTGCTCTGCAACTTTTCTATCTGAGCGCATCTATCAACCA 60
QY 1208 ATCTCTTACAACTCTATTTTCAAGAAGTACAGAGCGCGCTTTAACTGTGCTCGCA 1267
DB 61 ATCTCTTACAACTCTATTTTCAAGAAGTACAGAGCGCGCTTTTAACTGTGCTCGCA 120
QY 1268 AGGAAGTCCAGGCGCGAGAGCTTCCACAGAGCAGGACACTCGCGGGGGAAGTTGCAGGG 1327


```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-077-674-9

Query Match      16.5%; Score 228.8; DB 4; Length 1122;
Best Local Similarity 57.9%; Pred. No. 2.3e-37;
Matches 424; Conservative 0; Mismatches 307; Indels 1; Gaps 1;

QY   93 GCCTTTCCCTGGGCGGTGTTGCCGGTGACCGCTGTGCCTGTGCCTGTGCCTTTCGTCGT 152
Db   360 GCTCTTCCC CGCGCGCTGTCTGGCGGCGCTCACAGCCACCCTTCTTCGTGGT 419

QY   153 CGGGTTCAGCGGCAACTGGTGACCGGTATGCTCATCGGCGCTACCGGGACATGCGGAC 212
Db   420 GGGTATCGCTGGCAACTGCTTCACCATGCTGTGGTGTGGCGCTTCGCGAGCTCCGCAC 479

QY   213 CACCACAACTTGTACCTGGGCAGCATTGGCGCTGCCAGCCTACTCATCTCTCTCGGGCT 272
Db   480 CACCACAACTCTACCTGTCTCCAGCATGGCTTCTCCGATCTGCTCATCTTCCTCTGCAT 539

QY   273 GCGGTTGCAGCTGTACCGCTCTGGCGCTTCGGCGCTTCGGGCGCTGCTGCTG 332
Db   540 GCCCTTGAGCTCGTTCCGCTCTGGCAGTAGCGGCGCTTGGAACTTCGCGACCTCTCTG 599

QY   333 CCGGCTGTCCCTCTACCTGGGCGAGGCTGCACCTAGCCACGCTGTCGACATGACCGC 392
Db   600 CAAACTCTTCAAATCTGTAGTGAGAGCTGCACCTAGCCACGGTGTCTACCATCA CAGC 659

QY   393 GCTCAGCGTTCGAGCGCTACCTGGCCATCTGCGCGCGCTCTCGCGCCCGCGCTTGGTCA 452
Db   660 GCTGAGCTTCGAGCGCTACTTCGCCATCTGCTTCCACTTCGGGCCAAGGTGGTGCAC 719

QY   453 CCGGCGCGCGTTCGCGCGCTCATCGTGTGCTCTGGGCGCTGCGGCTGCTCTCTGCGCG 512
Db   720 CAAGGCGCGGTGAAGCTGTGTCATCTTCTCATCTGGGCGCTGGGCTTCTG CAGCGCGG 779

QY   513 TCCCTTCTGTTCTGTGGGCGTCGAGCAGGACCCCGGCATCT- CCGTAGTCCCGGCCC 571
Db   780 GCCCATCTTCTGTGTAGTCGGGGTGGAGCAGAGAACCGCACCCGACCTTTGGGACACCAA 839

QY   572 TCAATGCAACCGCGCGGATCGCCTCTCGCTCTGCGCTCTGTCGCCGCTCTCTGGGCTCT 631
Db   840 CGAGTGGCGCCCCACCGAGTTTGGGTGCGCTCTGGACTGCTCTACGGT CATGGTGGGT 899

QY   632 CGGGGCGCCACCGCGCTTCCCGCGCTGCGGGGCCGAGACCGCGGAGGCGCGCGCGCTGT 691
Db   900 GTCCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 959

QY   692 TCAGCCGGAATGCGGCGCGNAGCCCCCGGCGAGCTGGGGCGGCTGCTGTCATGCTGTGGG 751
Db   960 CAGGAAGCTGTGGCGGAGGAGGCGCGGAGTGTGTCTGTGGTGTCCCTCGCTCAGGGACCA 1019

QY   752 TCACACCGGCTACTTCTTCTCGCCCTTTCTGTGCTCTCAGCATCTCTCAGGGCTCATCG 811
Db   1020 GNACCACAAGCAACCGTGAAAATGCTGGGTGGTCTCAGCGCGCGCTCAGGCTTCTCT 1079

QY   812 GCGGCGAGCTGT 823
```

```

Db      1080  CGCGGGTCCTAT 1091
||||| |||
RESULT 11
US-09-077-675A-15
; Sequence 15, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Pleeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-15

Query March 16.4%; Score 228.4; DB
Best Local Similarity 68.4%; Pred.No.2.7e-3
Matches 316; Conservative 0; Mismatches 1

Qy      92  CGCCCTTTCCCTCGGGGCGCTGGTGGCGCGGTGCACCG
|||||
Db      104  GCCTGTTCCCGCGCTCCGCTGTCTGGCAGGCGTCACCG
|||||
Qy      152  TCGGGGTGAGCGGCAACGTTGGTGACCGTGAATGCTGA
|||||
Db      164  TGGGCATCTCAGGCAACCTGTCTCACTATGCTCGTGC
|||||
Qy      212  CCACCAACAATTGTACTCTGGCAGCATGGCCGTGT
|||||
Db      224  CCACCAACAACCTCTACCTGTCCAGCATGGCCCTTCTCT
|||||
Qy      272  TGCCGTTGCACCTGTACCGCCTCTCGCGGCTCGCGG
|||||
Db      284  TGCCGCTGGACCTGTCGCGCCTCTGGCAGTACGGG
|||||
Qy      332  GCCGCCTGTCCCTCTTACGTGGCGAGGGCTGCACCT
|||||

```

Db 344 GCAAACTCTTCCAGTGTGTGTCAGCGAGAGCTGCACCTACGCCACGGCTCTCACCACATCACCG 403
Qy 392 CGCTAGCGTCGAGCGCTACCTGGCCATCTGCCCGCGCTCCCGCGCGGTCTTGGTCA 451
Db 404 CGCTAGCGTCGAGCGCTACTTCGCCCATCTGCTTCCCTCTGCGGGCCAAAGGTGGTGGTCA 463
Qy 452 CCGCGCGCGCGCTCCGCGCTCATCGCTGTGCTCTGCGCGCTGGCGCTGCTCTCTGCGG 511
Db 464 CTAAGGCGCGCTGGAAGCTGGTCACTCTGTGTCACTGCGCGCGTGGCTTCTGCGAGCGGG 523
Qy 512 GTCCTCTTCTGTTCTCTGTTGGTGGCGTGCAGCAGACCCCGGCA 553
Db 524 GGCCCATCTTCGTGCTGGTGGCGTGGAGCAGCAACGCA 565

RESULT 12
US-09-077-674-15
; Sequence 15, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-674-15
Query Match 16.4%; Score 228.4; DB 4; Length 1092;
Best Local Similarity 68.4%; Pred. No. 2.7e-37;
Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
Qy 92 CGCCCTTTCCCTGGGGCGCTGGTGCCTGAGCCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 151
Db 104 CGCTGTTCCCGCTCCGCTGGCAGGCGTCAACGCCACCTCGGTGGCGCTCTTGTGGTGG 163

Qy 152 TCGGGGTGAGCGCGCAACGTGTGTGACCGTGTATCGGTGCTACCGGGACATCGGA 211
Db 164 TGGGCATCTCAGGCAACCTGCTCACTATGCTGGTGTGTCGCCGCTTCCGCGAGCTGGCA 223
Qy 212 CCACCAACCACTTGTACTTGGGCGAGCATGGCGGTGTCGCACTTACTCATCTGCTCGGGC 271
Db 224 CCACCAACCACTTACTTGTGTCAGCATGGCGCTTCTCGGATCTGCTCATCTTCTCTGTGCA 283
Qy 272 TGGCGTTGACCTGTACCGCTCTGCGGTCTGCGGCGCTTGGGTGTTTCGGGCGCTGCTCT 331
Db 284 TGGCGTGGACCTCGTCCGCTCTGGCAGTACCGGCGCTTGGAACTTCGGCGACCTGCTCT 343
Qy 332 GCGCGCTGCTCTACGTGGGCGAGGCTGCACCTACGCACTGCTGCTGCTGCTGCTGCTGCTG 391
Db 344 GCAAACTCTTCAGTTTGTGAGGAGCTGCACCTACGCCAGCTCTCACCACATCACCG 403
Qy 392 CGCTAGCGTCGAGCGCTACCTGGCCATCTGCCCGCGCTCCCGCGCGGTCTTGGTCA 451
Db 404 CGCTAGCGTCGAGCGCTACTTCCGCACTCTGCTTCCCTCTGCGGGCCAAAGGTGGTGGTCA 463
Qy 452 CCGGGCGCGCGTCCGCGCTCATCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
Db 464 CTAAGGCGCGCGTGAAGCTGGTCACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523
Qy 512 GTCCCTTCTTGTTCCTGTTGGGCGTGCAGCAGACCCCGGCA 553
Db 524 GGCCCATCTTCGTGCTGGTGGCGTGGAGCAGCAACGCA 565

RESULT 13
US-09-077-675A-14
; Sequence 14, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3129 base pairs
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-077-675A-14

Query Match 16.4%; Score 228.4; DB 3; Length 3129;
Best Local Similarity 68.4%; Pred. No. 3:2e-37;
Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 92 CGCCCTTTCCCTGGGGCGCTGTCGCGGTGACCGCTGTGCTGTGCTGTGCTGTGCTGCG 151
Db 104 CGCTGTTCCCGCTCCGCTGCTGGCAGCGCTCACCAGCACTGGTGGCGCTCTTCTGTGG 163
Qy 152 TCGGGGTGAGCGCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 211
Db 164 TGGGCATCTCAGGCAACCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 223
Qy 212 CCACCAACCACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 271
Db 224 CCACCAACCACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
Qy 272 TGGCGTTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
Db 284 TGGCGTTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343
Qy 332 GCGCGCTGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391
Db 344 GCAAACTCTTCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 403
Qy 392 CGCTCAGCGTCGAGCGCTACCTGCGCATCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 451
Db 404 CGCTGAGCGTCGAGCGCTACCTGCGCATCTGCGCGCATCTGCGCGCATCTGCGCGCA 463
Qy 452 CCGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCG 511
Db 464 CTAAAGGCGCGCGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
Qy 512 GTCCCTTCTGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
Db 524 GGCCCATCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565

RESULT 14
US-09-077-674-14
; Sequence 14, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-077-674-14

Query Match 16.4%; Score 228.4; DB 4; Length 3129;
Best Local Similarity 68.4%; Pred. No. 3:2e-37;
Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 92 CGCCCTTTCCCTGGGGCGCTGTCGCGGTGACCGCTGTGCTGTGCTGTGCTGTGCTGCG 151
Db 104 CGCTGTTCCCGCTCCGCTGCTGGCAGCGCTCACCAGCACTGGTGGCGCTCTTCTGTGG 163
Qy 152 TCGGGGTGAGCGCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 211
Db 164 TGGGCATCTCAGGCAACCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 223
Qy 212 CCACCAACCACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 271
Db 224 CCACCAACCACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
Qy 272 TGGCGTTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
Db 284 TGGCGTTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343
Qy 332 GCGCGCTGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391
Db 344 GCAAACTCTTCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 403
Qy 392 CGCTCAGCGTCGAGCGCTACCTGCGCATCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 451
Db 404 CGCTGAGCGTCGAGCGCTACCTGCGCATCTGCGCGCATCTGCGCGCATCTGCGCGCA 463
Qy 452 CCGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCG 511
Db 464 CTAAAGGCGCGCGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
Qy 512 GTCCCTTCTGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
Db 524 GGCCCATCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565

RESULT 15
US-09-077-675A-6
; Sequence 6, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 688.92 Seconds
(without alignments)
6983.649 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atgggcagcccttggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1078	77.6	1239	13	US-10-206-677-1 Sequence 1, Appli
2	1078	77.6	1239	15	US-10-225-567A-472 Sequence 472, App
3	1078	77.6	1239	15	US-10-290-078-13 Sequence 13, Appl
4	1078	77.6	1239	15	US-10-290-078-14 Sequence 14, Appl
5	1047.2	75.3	2040	12	US-10-417-820A-129 Sequence 129, App
6	1040.8	74.9	2040	12	US-10-417-820A-151 Sequence 151, App
7	234.8	16.9	1063	13	US-10-303-204A-1 Sequence 1, Appli
8	230	16.5	1029	13	US-10-303-204A-4 Sequence 4, Appli
9	228.8	16.5	870	15	US-10-225-567A-139 Sequence 139, App
10	228.8	16.5	1122	13	US-10-303-204A-9 Sequence 9, Appli
11	228.4	16.4	1092	13	US-10-303-204A-15 Sequence 15, Appl
12	228.4	16.4	3129	13	US-10-303-204A-14 Sequence 14, Appl
13	227.8	16.4	1088	13	US-10-303-204A-6 Sequence 6, Appli
14	227.8	16.4	1101	13	US-10-276-392-22 Sequence 22, Appl
15	227.8	16.4	1101	15	US-10-251-385-87 Sequence 87, Appli

16	227.8	16.4	1101	15	US-10-251-385-209	Sequence 209, App
17	153	11.0	836	13	US-10-303-204A-11	Sequence 11, Appl
18	134.8	9.7	801	11	US-09-782-974C-17	Sequence 17, Appl
19	134.8	9.7	1239	15	US-10-225-567A-556	Sequence 556, App
20	134.8	9.7	1248	12	US-10-417-820A-11	Sequence 11, Appl
21	134.8	9.7	1248	13	US-10-272-983-11	Sequence 11, Appl
22	134.8	9.7	1248	13	US-10-393-807-11	Sequence 11, Appl
23	127	9.1	1233	13	US-09-826-509-536	Sequence 536, App
24	127	9.1	1569	15	US-10-225-567A-431	Sequence 431, App
25	123	8.8	1257	13	US-09-826-509-534	Sequence 534, App
26	123	8.8	4131	13	US-10-101-510-417	Sequence 417, App
27	123	8.8	4131	13	US-10-101-510-752	Sequence 752, App
28	123	8.8	4131	15	US-10-225-567A-206	Sequence 206, App
29	123	8.8	4144	13	US-10-115-831-149	Sequence 149, App
30	122.2	8.8	1529	13	US-10-205-219-22	Sequence 22, Appl
31	122	8.8	1095	13	US-09-826-509-576	Sequence 576, App
32	122	8.8	1095	15	US-10-225-567A-319	Sequence 319, App
33	122	8.8	1285	15	US-10-265-872-2	Sequence 2, Appli
34	120	8.6	1349	12	US-10-240-145-96	Sequence 96, Appl
35	120	8.6	1535	12	US-10-240-145-10	Sequence 10, Appl
36	120	8.6	1535	15	US-10-146-419-12	Sequence 12, Appl
37	120	8.6	1535	15	US-10-146-123-12	Sequence 12, Appl
38	119	8.6	1212	13	US-10-353-690-9	Sequence 9, Appli
39	119	8.6	1212	15	US-10-083-188-13	Sequence 13, Appl
40	119	8.6	1212	15	US-10-083-188-82	Sequence 82, Appl
41	119	8.6	1212	15	US-10-251-385-113	Sequence 113, App
42	119	8.6	1212	15	US-10-251-385-223	Sequence 223, App
43	119	8.6	1212	15	US-10-225-567A-539	Sequence 539, App
44	119	8.6	1212	15	US-10-290-078-16	Sequence 16, Appl
45	119	8.6	1212	15	US-10-290-078-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-10-206-677-1
; Sequence 1, Application US/10206677
; Publication No. US20030186336A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences, Inc.
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; APPLICANT: Kulanter, Bruce G.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS RELATED TO
; TITLE OF INVENTION: GPCR 38, A G PROTEIN-COUPLED RECEPTOR (GPCR)
; FILE REFERENCE: 1920-1-8
; CURRENT APPLICATION NUMBER: US/10/206,677
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/250,251
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,452
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-677-1

Query Match	77.6%;	Score 1078;	DB 13;	Length 1239;
Best Local Similarity	89.1%;	Pred. No. 5.8e-255;		
Matches 1239;	Conservative	0;	Mismatches	0;
			Indels	151;
			Gaps	1;
Qy	1	ATGGGAGAGCCCTTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGCGTGGCCCC	60	
Db	1	ATGGGAGAGCCCTTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGCGTGGCCCC	60	
Qy	61	GCCTCCGCGCTTGCACGAGCGCCGCTGCTCCCTTTCCCTGGGGGGCGTGTGTCGC	120	
Db	61	GCCTCCGCGCTTGCACGAGCGCCGCTGCTCCCTTTCCCTGGGGGGCGTGTGTCGC	120	

Db 421 TGCGCCGCTCCGCGCCGCTTTGGTACCCGCGCGCGCTCCGCGGCTCATCGCT 480
Qy 481 GTGCTCTGGGCGGTGGCGCTGCTCTCTGCGCGTCCCTTCTTGTCTCTGGTGGGCGTGGAG 540
Db 481 GTGCTCTGGGCGGTGGCGCTGCTCTCTGCGCGTCCCTTCTTGTCTCTGGTGGGCGTGGAG 540
Qy 541 CAGAACCCCGGATCTCCGTAGTCCCGGCTCAATGCGACCGCGCGATCGCTCTCTCG 600
Db 541 CAGAACCCCGGATCTCCGTAGTCCCGGCTCAATGCGACCGCGCGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCTGCGCGCTCTCTGCGCTCTGCGCGCGCGACCGCGTCCCGCGCTCG 660
Db 601 CCTCTCGCTCTGCGCGCTCTCTGCGCTCTGCGCGCGCGACCGCGTCCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATCCGCGCGAGCGCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATCCGCGCGAGCGCGCG 720
Qy 721 CAGTGGGCGCTGCGTGTGATCTGCTGTGGGTCAACACCGCTACTTCTTCTGCGCTTT 780
Db 721 CAGTGGGCGCTGCGTGTGATCTGCTGTGGGTCAACACCGCTACTTCTTCTGCGCTTT 780
Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCGGAGCTGTGGAGAGCGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCGGAGCTGTGGAGAGCGCGCGCG 840
Qy 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGCGCGCACCGGAGAGCGCTCGCGTCTG 900
Db 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCGCACCGGAGAGCGCTCGCGTCTG 900
Qy 901 CGTAAGTGGAGCGCGCGTGGTTCAGAGAGCGCTGCTGAGTCCGCGCGCGCGGAGC 960
Db 901 C----- 901
Qy 961 GCGCAAGCGTGGGTCCCTTCCCTTCCCTGCTGCCCGAGCTCTGGGCGCGCTTCCAGTCCC 1020
Db 902 ----- 902
Qy 1021 TTTCTATTTCGATTCCAGCTCCACCGCGTGGTGGTCTTGGCAITTAATAATTG 1080
Db 902 -----TGGTGGTGGTCTTGGCAITTAATAATTG 929
Qy 1081 CTGTTGCTCCCTTCCAGTGGCAGAACTATTACATAAACAAGAGATTTCGCGGATGAT 1140
Db 930 CTGTTGCTCCCTTCCAGTGGCAGAACTATTACATAAACAAGAGATTTCGCGGATGAT 989
Qy 1141 GTACTTCTCAGTACTTTAACTGCTGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
Db 990 GTACTTCTCAGTACTTTAACTGCTGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
Qy 1201 CAACCCAACTCTCTACAACTTCTTCAAGAAGTACAGAGCGCGGCTTTAACTGCT 1260
Db 1050 CAACCCAACTCTCTACAACTTCTTCAAGAAGTACAGAGCGCGGCTTTAACTGCT 1109
Qy 1261 GCTCGAAGGAGTCCAGGCGGAGAGGCTTCCAGAGAGAGGAGACTCGCGGGGAGT 1320
Db 1110 GCTCGAAGGAGTCCAGGCGGAGAGGCTTCCAGAGAGAGGAGACTCGCGGGGAGT 1169
Qy 1321 TGAGGGGACACTGGAGAGAGACAGTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1380
Db 1170 TGAGGGGACACTGGAGAGAGACAGTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

RESULT 3
US-10-290-078-13
; Sequence 13, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:

; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; FILE OF INVENTION: 14395, 14618, 17692 or 58874
; FILE REFERENCE: MPI2001-288P1(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-290-078-13

Query Match 77.6%; Score 1078; DB 15; Length 1239;
Best Local Similarity 89.1%; Pred. No. 5.8e-255;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

Qy 1 ATGGGAGAGCCCTTGGAAACGAGACGCGCCCGAGGGGCGCGGAGCGCGTGGCCCC 60
Db 1 ATGGGAGAGCCCTTGGAAACGAGACGCGCCCGAGGGGCGCGGAGCGCGTGGCCCC 60
Qy 61 GCGCTGCGCGCTTGGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGCGCTGGTGGCG 120
Db 61 GCGCTGCGCGCTTGGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGCGCTGGTGGCG 120
Qy 121 GTACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTTACCGGGAGATCGGGAGACCAACCACTTGTACTCTGGGCGAGCATG 240
Db 181 ATGCTGATCGGGCGCTTACCGGGAGATCGGGAGACCAACCACTTGTACTCTGGGCGAGCATG 240
Qy 241 GCGGTGTGCGACCTTACTATCTGCTGCGGCTGCGGTTCGACCTGTACCGGCTCTGGCGC 300
Db 241 GCGGTGTGCGACCTTACTATCTGCTGCGGCTGCGGTTCGACCTGTACCGGCTCTGGCGC 300
Qy 301 TCGCGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 360
Db 301 TCGCGCGCTTGGGTGTTGGGGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 360
Qy 361 TGACACCTTACCGGAGCTGCTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Db 361 TGACACCTTACCGGAGCTGCTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Qy 421 TGCGCGCGCTTCCGCGCGCTGCTTGGGTCAACCGCGCGCGCTGCTGCGCGCGCTCATCGCT 480
Db 421 TGCGCGCGCTTCCGCGCGCTGCTTGGGTCAACCGCGCGCGCTGCTGCGCGCGCTCATCGCT 480
Qy 481 GTGCTCTGGGCGCTGGGCGCTGCTCTGCGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTCTGGGCGCTGGGCGCTGCTCTGCGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGAACCCCGGATCTCCGTAGTCCCGGCTCAATGGACCGCGGATCGCTCTCTCG 600
Db 541 CAGAACCCCGGATCTCCGTAGTCCCGGCTCAATGGACCGCGGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCTGCGCGCTCTCTGCGCTCTGCGCGCGCGACCGCGTCCCGCGCTCG 660
Db 601 CCTCTCGCTCTGCGCGCTCTCTGCGCTCTGCGCGCGCGACCGCGTCCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATGCGCGCGAGCGCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATGCGCGCGAGCGCGCG 720
Qy 721 CAGTGGGCGCTGCGTGTGATCTGCTGTGGGTCAACACCGCTACTTCTTCTGCGCTTT 780
Db 721 CAGTGGGCGCTGCGTGTGATCTGCTGTGGGTCAACACCGCTACTTCTTCTGCGCTTT 780
Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCGGAGCTGTGGAGAGCGCGCGCG 840

```

Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGACCGCGCGCG 840
Qy 841 CTGAGGCGCGCGCTCGGGCGGAGAGAGGCGCACCGGAGACCGCTCGCGCTCTG 900
Db 841 CTGAGGCGCGCGCGCTCGGGCGGAGAGAGGCGCACCGGAGACCGCTCGCGCTCTG 900
Qy 901 CGTAAGTGGAGCGCGCTGTGTTCCAAAGACCGCTGTGCTGAGTCCGCGCGCGGAGAC 960
Db 901 C----- 901
Qy 961 GCGCAACGCTGGGTCCCTTCCCTCTGCTGCCAGCTCTGGGGCGCGCTTCCAGCTCCC 1020
Db 902 ----- 901
Qy 1021 TTTCTATTTCGATTCAGGCTCCAGCGCTCCAGCGCGGTCTGTGSCATTTATAATTG 1080
Db 902 -----TGGTGGTGGTCTTGGCATTTATAATTG 929
Qy 1081 CTGGTTGCCCTTCCAGCTTGGCAGAAATCAITTTACATAAACAAGAGATTCGGGGATGAT 1140
Db 930 CTGGTTGCCCTTCCAGCTTGGCAGAAATCAITTTACATAAACAAGAGATTCGGGGATGAT 989
Qy 1141 GTACTTCTCAGTACATTTAAACATGCTGCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
Db 990 GTACTTCTCAGTACATTTAAACATGCTGCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
Qy 1201 CAACCCAACTCTTACAACTCTTTTCAAGAGAGTACAGAGCGCGGCTTTAACTGCT 1260
Db 1050 CAACCCAACTCTTACAACTCTTTTCAAGAGAGTACAGAGCGCGGCTTTAACTGCT 1109
Qy 1261 GCTCGCAAGGAAGTCCAGGCGGAGAGCTTCCACAGAGAGGAGACACTCGCGGGGAAGT 1320
Db 1110 GCTCGCAAGGAAGTCCAGGCGGAGAGCTTCCACAGAGAGGAGACACTCGCGGGGAAGT 1169
Qy 1321 TGAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCGCTAACGTGAAGAC 1380
Db 1170 TGAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCGCTAACGTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

RESULT 4
US-10-290-078-14
; Sequence 14, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17692 or 58874
; FILE REFERENCE: MPI2001-288P1(N)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1239)
US-10-290-078-14
Query Match 77.6%; Score 1078; DB 15; Length 1239;
Best Local Similarity 89.1%; Pred. No. 5.8e-255;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

Qy 1 ATGGGACAGCCCTGGAAACGGCAGCGACGGCGCGCGGGGCGCGGGAGCGCGGTGGCCC 60
Db 1 ATGGGACAGCCCTGGAAACGGCAGCGACGGCGCGCGGGGCGCGGGAGCGCGGTGGCCC 60

```

```

Qy 61 GCGCTGCCCGCTTGGCAACGAGCGCGCTGTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
Db 61 GCGCTGCCCGCTTGGCAACGAGCGCGCTGTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
Qy 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Qy 181 ATGCTCATCGGGCGCTACCGGGACATGCGGACACACCAACAACTGTGTACCTGGGACGATG 240
Db 181 ATGCTCATCGGGCGCTACCGGGACATGCGGACACACCAACAACTGTGTACCTGGGACGATG 240
Qy 241 GCGGTGTGCGACCTACTCATCTGCTCGGGCTGCGGTTCGACCTGTACCGCTCTGGCGC 300
Db 241 GCGGTGTGCGACCTACTCATCTGCTCGGGCTGCGGTTCGACCTGTACCGCTCTGGCGC 300
Qy 301 TCGCGCCCTTGGGTGTTCGGGCGCTGCTGTGCGCGCTGTCTCTCTACGTGGGCGAGGGC 360
Db 301 TCGCGCCCTTGGGTGTTCGGGCGCTGCTGTGCGCGCTGTCTCTCTACGTGGGCGAGGGC 360
Qy 361 TGCACCTACGCCACGCTGTGACATGACCGCGCTCAGCGTTCGAGCGCTACCTGGCGATC 420
Db 361 TGCACCTACGCCACGCTGTGACATGACCGCGCTCAGCGTTCGAGCGCTACCTGGCGATC 420
Qy 421 TCGCGCGCTGTGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 480
Db 421 TCGCGCGCTGTGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 480
Qy 481 GTGCTCTGGGCGCTGTGCGCTGTGCTGTGCGCGCTGTCTCTCTGTGCTGTGCTGTGCTGTG 540
Db 481 GTGCTCTGGGCGCTGTGCGCTGTGCTGTGCGCGCTGTCTCTCTGTGCTGTGCTGTGCTGTG 540
Qy 541 CAGGACCGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600
Db 541 CAGGACCGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600
Qy 601 CTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 660
Db 601 CTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 720
Qy 721 CAGCTGGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 780
Db 721 CAGCTGGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 780
Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGACCGCGCGGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGACCGCGCGGCGCG 840
Qy 841 CTGCGAGGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 900
Db 841 CTGCGAGGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 900
Qy 901 CGTAACTGGAGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 960
Db 901 C----- 901
Qy 961 GCGCAACGCTGGGTCCCTTCCCTCTGCTGCCAGCTCTGGGGCGCGCTTCCAGCTCCC 1020
Db 902 ----- 901
Qy 1021 TTTCTATTTCGATTCAGGCTCCAGCGCTCCAGCGCGGTCTGTGSCATTTATAATTG 1080
Db 902 -----TGGTGGTGGTCTTGGCATTTATAATTG 929
Qy 1081 CTGGTTGCCCTTCCAGCTTGGCAGAAATCAITTTACATAAACAAGAGATTCGGGGATGAT 1140
Db 930 CTGGTTGCCCTTCCAGCTTGGCAGAAATCAITTTACATAAACAAGAGATTCGGGGATGAT 989

```


153 CGGGTGAGCGGCAACGTGGTGTGACCGTGATGTCGGGCGCTACGGGACATGGGAC 212

168 GGGTATCGTGGCAACCTGTCAACCATGTGTGGTGGTGTGCGCTTCGCGGAGCTGGCAC 227

```

RESULT 15
US-10-251-385-87
; Sequence 87, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 87
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-87

```

```

; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 87
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-87

Query Match          16.4%;   Score 227.8;   DB 15;   Length 1101;
Best Local Similarity 68.3%;   Pred. No. 2.9e-46;
Matches 316;   Conservative 0;   Mismatches 147;   Indels 0;   Gaps 0;

Qy      93  GCCTTTCCCTGGGGGGCGCTGTGTGCGGTGACCGGTGTGTGTGCTGTGTGCTGTTCGTTCGT 152
          |||||
Db      108  GTCTTCCCGCGCGCGCTGTGTGCGGGCGGTACAGCCACTGCGTGGCACTTTCGTGTT 167
          |||||

```

Qy	153	CGGGTGAGCGGCAACGTGTGACCGGTGATCGTGGCGGCTACCGGACATGGGAC	212
Db	168	GGGTATCGTGGCAACCTGTCAACATGCTGGTGGTGTGCGGCTTCGCGAGCTGGGCAC	227
Qy	213	CACCACCAACTTGTACCTGGGAGCATGGCGGTGTCGACCTACTCATCTGCTCGGGCT	272
Db	228	CACCACCAACTCTACTGTCCAGCATGGCTTCTCCGATCTGCTCATCTTCTCTGCA	287
Qy	273	GCGTTGACCTGTACCGCCTCTGGCGCTCGCGCCCTGGGTGTTCGGGCCGCTGCTCTG	332
Db	288	GCCCTGGACCTCGTTGCGCTCTGGCAGTACCGGCCCTGGAACTTCGGCGACCTCCTCTG	347
Qy	333	CGGCTGTCCCTCTACGTGGGAGGGCTGCACCTAGCCACCGTGTGCACATGACCGC	392
Db	348	CAAACTCTTCCAATTCTGTCAGTGAGAGTGCACCTACGCCACGGTGTCAACCATCACGC	407
Qy	393	GCTCAGGTCGAGCGCTACCTGGCCATCTCGCGCGCTCGCGCGCGCTTGTGGTCAC	452
Db	408	GCTGAGGCTCGAGCGCTACTTCGCCATCTGCTTCCCACTCGGGCCCAAGGTGGTGGTCAC	467
Qy	453	CCGGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTGGGCGGTGGCGGTGCTCTGCGCG	512
Db	468	CAAGGGCGGGTGAAGCTGGTCATCTTCGTCACTGGGCGGTGGCTTCTGCAGCGCGG	527
Qy	513	TCCCTTCTTGTTCGTGGGCGTGCAGAGGACCCCGGCATC	555
Db	528	GCCCATCTTCGTGCTAGTCGGGGTGGAGCACGAGACGGCACC	570

Search completed: January 1, 2004, 06:44:59
 Job time : 693.92 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:38:58 ; Search time 3259.02 Seconds
(without alignments)
10366.072 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atgggcagccctggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	230	16.5	4435	11 AK049671	AK049671 Mus muscu
2	226.8	16.3	500	10 BF603623	BF603623 269181 MA
3	219.2	15.8	608	14 BY724644	BY724644 BY724644
4	218	15.7	590	10 BF513101	BF513101 UI-H-BW1-

5	186.4	13.4	843	13	BU553576	BU553576 AGENCOURT
6	173.4	12.5	849	13	BU568940	BU568940 AGENCOURT
7	141.2	10.2	504	28	AZ451922	AZ451922 1M0251C06
8	127	9.1	855	12	BI757121	BI757121 603030882
9	127	9.1	1010	12	BM925480	BM925480 AGENCOURT
10	125.4	9.0	634	10	EG714306	EG714306 602669812
11	125.4	9.0	1055	12	BM808829	BM808829 AGENCOURT
12	125.2	9.0	540	12	BM737726	BM737726 K-EST0000
13	124.8	9.0	1027	12	BM543497	BM543497 AGENCOURT
14	124	8.9	1144	12	BM546496	BM546496 AGENCOURT
15	122.8	8.8	941	29	CNS036WH	AL230570 Tetraodon
16	119	8.6	1603	12	BM808823	BM808823 AGENCOURT
17	117.8	8.5	624	9	AU244292	AU244292 AU244292
18	111.8	8.0	738	12	BI600170	BI600170 603246569
19	111.4	8.0	426	14	CB695276	CB695276 AMGNNUC:N
20	111.4	8.0	574	14	CB608597	CB608597 AMGNNUC:N
21	111.2	8.0	928	12	BI758259	BI758259 603029910
22	109.6	7.9	993	13	BQ721315	BQ721315 AGENCOURT
23	109.2	7.9	448	10	BF323227	BF323227 maa38a07.
24	106.4	7.7	3277	11	AK036756	AK036756 Mus muscu
25	105.6	7.6	2760	11	AK081073	AK081073 Mus muscu
26	105	7.6	427	13	BY284735	BY284735 BY284735
27	104.8	7.5	828	12	BI597845	BI597845 603245326
28	104.6	7.5	660	14	BY728041	BY728041 BY728041
29	103.6	7.5	866	14	CD246184	CD246184 AGENCOURT
30	102.6	7.4	980	12	BM543468	BM543468 AGENCOURT
31	102.4	7.4	419	13	BY273314	BY273314 BY273314
32	102.2	7.4	994	14	BY705540	BY705540 BY705540
33	102.2	7.4	1189	11	AK005368	AK005368 Mus muscu
34	99.4	7.2	1233	14	CA975828	CA975828 AGENCOURT
35	99	7.1	397	14	CB772331	CB772331 AMGNNUC:N
36	98.8	7.1	1918	11	AK053776	AK053776 Mus muscu
37	98.4	7.1	751	14	CB154463	CB154463 K-EST0212
38	97.8	7.0	982	13	BX415111	BX415111 BX415111
39	97	7.0	785	12	BI754749	BI754749 603025465
40	97	7.0	843	29	CNS048GJ	AL305164 Tetraodon
41	96.8	7.0	670	29	AG046172	AG046172 Pan trogl
42	96.8	7.0	720	12	BI753905	BI753905 603022907
43	96.8	7.0	1307	11	CNSLT11BD	BX248780 human ful
44	96.6	6.9	555	10	BE751626	BE751626 203809 MA
45	96.6	6.9	825	9	AU079556	AU079556 AU079556

ALIGNMENTS

RESULT 1:
AK049671
LOCUS
DEFINITION
Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length enriched library, clone:C530020I22 product:GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 homolog [Rattus norvegicus], full insert sequence.
AK049671
VERSION
AK049671.1 GI:26340405
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
2
99279253
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374


```

Research 6, 791-806.
TAG_Seq=None found"
BASE COUNT      130 a      173 c      157 g      130 t
ORIGIN

Query Match      15.7%; Score 218; DB 10; Length 590;
Best Local Similarity 100.0%; Pred. No. 3.6e-35;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 835 CGGCGCTGCGAGCGCGCGCGCTCGGGGCGGAGAGAGCCACCGGAGAGCGTCCGC 894
DB 1 CGGCGCTGCGAGCGCGCGCGCTCGGGGCGGAGAGAGCGGAGAGCGTCCGC 60
QY 895 GTCTCGGTAAAGTGAGCGCGCGGTTCAGAGCGCTGCTGCAAGTCCGCGCGCG 954
DB 61 GTCTCGGTAAAGTGAGCGCGCGGTTCAGAGCGCTGCTGCAAGTCCGCGCGCG 120
QY 955 GGGACCGCGCAAGCGTGGGTCCCTTCCCTGCTGCGCGCGCTGCGCGCGCTTCCA 1014
DB 121 GGGACCGCGCAAGCGTGGGTCCCTTCCCTGCTGCGCGCGCTGCGCGCGCTTCCA 180
QY 1015 GCTCCCTTCTCTATTTTCGATTTCAGCTTCCAGCTTCCACCGCGCG 1052
DB 181 GCTCCCTTCTCTATTTTCGATTTCAGCTTCCAGCTTCCACCGCGCG 218

RESULT 5
BU553576      843 bp      mRNA      linear      EST 16-SEP-2002
LOCUS
DEFINITION
AGENCOURT_10242213 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6577973 5', mRNA sequence.
ACCESSION
BU553576      1 GI:22903848
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 843)
AUTHORS
NTH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2778 row: p column: 05
High quality sequence stop: 534.
FEATURES
source
location/Qualifiers
1. 843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6577973"
/tissue_types="teratocarcinoma, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      213 a      169 c      212 g      247 t      2 others
ORIGIN

Query Match      13.4%; Score 186.4; DB 13; Length 843;
Best Local Similarity 91.0%; Pred. No. 1.5e-28;
Matches 243; Conservative 0; Mismatches 17; Indels 7; Gaps 4;

QY 1057 GGTGGTTCTGGCATTATTAATTTGCTGGTTGCCCTTCCACGTTGGCAGATCATTTACAT 1116
DB 534 GGTGGTTCTGGCATTATTAATTTGCTGGTTGCCCTTCCACGTTGGCAGATCATTTACAT 593
QY 1117 AAACACGGAAGATTCCGGGATGATGACTTCTCTCAGTACTTTTAAACATCGTCCTCTGCA 1176
DB 594 AAACACGGAAGATTCCGGGATGATGACTTCTCTCAGTACTTTTAAACATCGTCCTCTGCA 653
QY 1177 ACTTTTCTATCGAGCGCATCTATCAACCAATCTCTCAACCTC-ATTTCAAAGAAGT 1235
DB 654 ACTTTTCTATCGAGCGCATCTATCAACCAATCTCTCAACCTCAATTTTCAAGAAGT 713
QY 1236 AC-AGAGCGCGCGCTTTAAA---CTGCTGCTCGCAAGGAAGTCCAGGCCGAG-AGGCT 1289
DB 714 ACAAGAGCGCGCGCTTTAAAAGTCTGCTTCCAAAGGAAGTCCAGGCCGAGAGGCT 773
QY 1290 TCCACAGAGACGAGGACACTGCGGGGG 1316
DB 774 CCACAGAGACGAGGACACTGCGGGGG 800

RESULT 6
BU568940/c
LOCUS
DEFINITION
AGENCOURT_10400272 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6616239
IMAGE:6616239 5', mRNA sequence.
ACCESSION
BU568940      1 GI:22919240
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 849)
AUTHORS
NTH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2859 row: j column: 15
High quality sequence stop: 429.
FEATURES
source
location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6616239"
/lab_hosts="DH10B (T1 phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctcgcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
BASE COUNT      199 a      229 c      193 g      228 t
ORIGIN

```

Query Match 12.5%; Score 173.4; DB 13; Length 849;
 Best Local Similarity 88.6%; Pred. No. 7.3e-26;
 Matches 233; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

Qy 1132 GCGGATGATCTCTCTCAGTACTTAA-CATCGTCGCTCG--CAACTTTTCTATCT 1188
 Db 698 GGAAGAGGAGTACTTTCTCAGTACTTTAACTTGTGCGTTTGGCAACTTTTTTTTGT 639

Qy 1189 GAGCGATCTATCAACCAATCTCTCTACAACTTCAAAAGAGTACAGAG-CGCGCG 1247
 Db 638 AGGGCATTTATCAACCAATCTTACACCTTCAAGAGTACAGAGCGCGCG 579

Qy 1248 CTTTAACTGCTGCTCGCAAGAGTCCAGGCGAGAGCTTCCACAGAGCAGGACA 1307
 Db 578 CTTTAACTGCTGCTCGCAAGAGTCCAGG-CGAGAGGCTTCCACAGAGCAGGACA 520

Qy 1308 CTGCGGGGAGTTCGAGGGGACACTGAGGAGACAGCTGGGCTACACCGACAGCG 1367
 Db 519 CTGCGGGGAGTTCGAGGGGACACTGAGGAGACAGCTGGGCTACACCGACAGCG 460

Qy 1368 CTAACGTGAAGACGATGGGATAA 1390
 Db 459 CTAACGTGAAGACGATGGGATAA 437

RESULT 7
 AZ451922/c 504 bp DNA linear GSS 04-OCT-2000
 LOCUS IM0251C06R Mouse 10kb plasmid UUGCM library Mus musculus genomic
 DEFINITION clone UUGCM0251C06 R, genomic survey sequence.
 ACCESSION AZ451922
 VERSION AZ451922.1 GI:10608187
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 504)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
 and Wright, D., Weiss, R.,
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0251 row: C column: 06
 Seq primer: CACACGAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 504.
 Location/Qualifiers
 1..504
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCM0251C06"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGCM library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 156 a 110 c 107 g 131 t
 ORIGIN

Query Match 10.2%; Score 141.2; DB 28; Length 504;
 Best Local Similarity 78.7%; Pred. No. 3.2e-19;
 Matches 181; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Qy 1052 GTGGTGGTCTGCGCATTTAATTGC-TGGTTGCCCTTCCAGTTGGCAGAAATCAT 1110
 Db 250 GTGGTGGAGGTTCTGCGCATTCAGTTGCTGGTTGCCCTTCCAAAGTTGGCAGAAATCTT 191

Qy 1111 TTACATAAAACACGGAAGATTCCGGGATGATGTTCTCTCAGTACTTTAAACATCGTCGC 1170
 Db 190 TTTCATGAACACTGAAATTCGGGGTAAATGCACTCTCTCGATATCTTAATATGGTCAC 131

Qy 1171 TCTGCAACTTTTCTATCTGAGGCATCTATCAACCAATCTCTCAACCTCTTTCAAA 1230
 Db 130 TCTCAACTTCTATACCGAGTGCATCCATCAAGCCCATCTCCCAACCTCTATGTCCA 71

Qy 1231 GAAGTACAGCGGGCGCTTTAACTGCTGCGCAAGGAAGTCCAGGC 1280
 Db 70 TTAGTACAAGACGCTGTCCAGAACTGCTTCCGCGAGACATGCCAGGC 21

RESULT 8
 BI757121 855 bp mRNA linear EST 25-SEP-2001
 LOCUS 603030882F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:520868 5',
 DEFINITION mRNA sequence.
 ACCESSION BI757121
 VERSION BI757121.1 GI:15748699
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 855)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11503 row: d column: 21
 High quality sequence stop: 803.
 Location/Qualifiers
 1..855
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:520868"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"

FEATURES
 source

/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH MGC Library."

BASE COUNT 103 a 300 c 299 g 153 t
ORIGIN

Query Match 9.1%; Score 127; DB 12; Length 855;
Best Local Similarity 55.2%; Pred. No. 3.2e-16;
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

QY 27 CGGCCCGAGGGGGGAGCGCGTGGCGCGCGCTCGCGCTTGGACGAGCGCG 86
DB 60 CAGACGCCGCGCGCGCGCGCGAGTCTCAACCCGGGCTGAGCTGGACCGCGCT 119
QY 87 CTGCTCGCCCTTTCCCTGGGGGGCTGTGTGCGGGTACCGCTGTGTGCTGTGTT 146
DB 120 GGCGCTGGACACTCGCTCTGGGCGAGGTGCTTTCACCGCGCTCTACGCACTCATCTG 179
QY 147 GTGCTGGGGGTGAGCGGCAACGTG-----GTGACCGTGATGCTGATCGGGCGCTACCG 200
DB 180 GGCGCTGGGCGCGCGGCAATGCGCTGCCGTGCACGTGGTGTGAAGGCGCGGCGCG 239
QY 201 GGACATGCGGACCAACACCACTGTGACCTGGGCGAGCATGGCGGTGTCGACCTACTCAT 260
DB 240 GCAGCGGGGGCGGCTGCGCGCACCACTGCTGACGCTGGCGCTCGGGGCTGTGTGCT 299
QY 261 CTGCTCGGGCTGCGGCTTGCACCTGTACCGCTCTGGCGCTCGCGG---CCCTGGGTGTT 317
DB 300 GCTGTTGGGCTGCGGCTGAGCTTACAGCTTGTGTGTTTCACTACCCCTTGGTCTT 359
QY 318 CGGGCGGCTGCTTGCGCGCTGCTCCCTTACGTGAGGCGAGGCTGACACTACCGCAAGCT 377
DB 360 CGCGGACCTGGGCTGCGCGGCTACTACTTCTGTGTGACAGCTGTGCGCTTACCGCAGGT 419
QY 378 GCTGCACATGACCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 437
DB 420 GCTGAGGCTGGAGGCTGAGCGCGAGCTGCTGCTGAGCTGCTGAGCGCTGCTGAGCTG 479
QY 438 CGCGCTCTTGGTACCGCGCGCGCTGCGCGCTCATCGCTGTGTGCTGTGGCGCTGGC 497
DB 480 CGCGAGCTGCTGAGCGCAAGCGGACCGGCTGGTGGTGGCTGTGCTGCTGAGCGCGCTC 539
QY 498 GCTGCTCTGCGGCTGCTTCTTGTTCCTGTTGGGCGTGCAGAGGACCGCG 550
DB 540 GCTCGGCTCGGCTGCGGCTGCGGCTCATCATGCGGCGAGAGCAGCAACTCG 592

RESULT 9
BM925480
LOCUS
DEFINITION 1010 bp mRNA linear EST 12-MAR-2002
AGENCY 6625013 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5763267
5' mRNA sequence.

ACCESSION BM925480
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1010)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12814 row: n column: 04
High quality sequence stop: 703.
Location/Qualifiers
1. .1010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5763267"
/lab_host="DH10B"
/clone_lib="NIH MGC 114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH MGC Library."

BASE COUNT 149 a 348 c 337 g 176 t
ORIGIN

Query Match 9.1%; Score 127; DB 12; Length 1010;
Best Local Similarity 55.2%; Pred. No. 3.2e-16;
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

QY 27 CGGCCCGAGGGGGGAGCGCGTGGCGCGCGCTCGCGCTTGGACGAGCGCGCG 86
DB 82 CAGACGCCGCGCGCGCGCGCGAGTCTCAACCCGGGCTGAGCTGGACCGCGCT 141
QY 87 CTGCTCGCCCTTTCCCTGGGGGGCTGTGTGCGGGTACCGCTGTGTGCTGTGCTGTGTT 146
DB 142 GGCGCTGGACACTCGCTCTGGGCGCAAGGTGCTTTCACCGCGCTCTACGCACTCATCTG 201
QY 147 GCTGCTGGGGGTGAGCGGCAACGTG-----GTGACCGTGATGCTGATCGGGCGCTACCG 200
DB 202 GGCGCTGGGCGCGCGGCAATGCGCTGCCGTGCACGTGGTGTGAGCGCGGCGCG 261
QY 201 GGACATGCGGACCAACCACTTGTACCTGGGCGAGCATGGCGCTTGGACCTACTCAT 260
DB 262 GCGCGGGGGCGCTGCGGCAACCACTGCTGAGCTGCGGCTGCTGCTGCTGCT 321
QY 261 CTGCTCGGGCTGCGGCTTTCAGCTGTGACCGCTCTGCGGCTGCGGCTGCGGCTGCT 317
DB 322 GCTGCTGGGCTGCGGCTGAGCTCTACAGCTTGTGTGCTTCCACTACCGCTGCGGCTT 381
QY 318 CGGGCGCTGCTGCGGCTGCTTCCCTTACGTGGGCGAGGGCTGCACCTACGCGCGCT 377
DB 382 CGCGACCTGGGCTGCGGCTTACTTCTGTCGACAGCTGTGCGCTTACGCGCGCT 441
QY 378 GCTGCACATGACCGCTCAGCGCTGAGCGCTACCTGGGCGATCTGCGCGCGCTGCGCG 437
DB 442 GCTGAGCTGGAGGCTGAGCGCGAGCGCTGCTGAGCGCTGAGCGCTGAGCGCTGAG 501
QY 438 CGCGCTTGTGGTACCGCGCGCGCTGCGGCTCATCGCTGTGTGCTGTGGCGCTGGC 497
DB 502 CGCGACCTGCTGAGCGCAACCGCGCGCGCTGGTGGTGGCTCTCTGTTGGCGCGCTC 561
QY 498 GCTGCTCTGCGGCTGCGCTTCTTGTTCCTGTTGGGCGTGCAGAGGACCGCG 550
DB 562 GCTCGGCTCGGCTCGGCTGCGGCTCATCATGCGGCGAGAGCAGCAACTCG 614

RESULT 10
BG714306
LOCUS
DEFINITION

BG714306 634 bp mRNA linear EST 08-MAY-2001
602669812F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792730 5',
mRNA sequence.


```

ACCESSION   BG714306
VERSION     BG714306.1  GI:13993237
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 634)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            cDNA Library Preparation: Mikhail Palkovits, M.D., Ph.D.
            Tissue Procurement: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10671 row: g column: 03
            High quality sequence stop: 633.

FEATURES             source
    1..634
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4792730"
        /tissue_type="hypothalamus"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_96"
        /notes="Organ: brain; Vector: pBluescriptR (modified
        pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
        ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
        size-selected for average insert size 2.3 kb and
        normalized to ROT 5. This is a primary library enriched
        for full-length clones and constructed using the
        Cap-trapper method (Carninci, in preparation). Library
        constructed by M. Brownstein (NIH/NHGRI, National
        Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT      74 a      221 c      243 g      96 t
ORIGIN
Query Match      9.0%; Score 125.4; DB 10; Length 634;
Best Local Similarity 55.0%; Pred. No. 6.4e-16;
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGGCGCGGAGCGCGCGTGGCCCGCGCTGCCGCTTGCAGCAGGCGCG 86
DB 49 CAGCAGCCCGCGCCCGCGCGCCAGCTCCAAACCCGGGCTGAGCTGACGCCCGGCT 108
QY 87 CTGCTGCGCTTTCCCTCGGGGCGCTGTGTGCGGTGACCGCTGTGTGCTGTGCTGTT 146
DB 109 GGGCGTGGACACTCGCTCTGTGGGCAAGTGTCTTCCCGCGCTCTACGCACTCATCTG 168
QY 147 CGTCTGCGGGTGTAGCGGCAACGTG-----GTGACCGTGATCTGATCGGGCGCTACCG 200
DB 169 GGGCGTGGGCGCGCGGGCAATGCGTGTCCGTGCACGTGTGTGAAGCGCGGGCGCG 228
QY 201 GGACATCGGACACCAACCACTTGATCTGTGGGCGAGCATGGCCGTGTGCCACCTACTCAT 260
DB 229 GCAGCGGGGCGCTGCGGCCACACAGTGTCTACGCTGGCGCTCGCGGGCTGTGTGCT 288
QY 261 CTGCTCGGCTGCGGTTCGACCTGTACCGCTCTGTGCGCTCGCG---CCCTGGGTGTT 317
DB 289 GCTGTGCGGCTGCGGGTGTGAGCTCTACAGTTCGTGTGTTCCACTACCGCTGGTCTT 348
QY 318 CGGGCGCTGCTGTGCGCGCTGTCCCTCTACGTGGGCGAGGGTGTACCTACCGCAGCT 377
DB 349 CGGCGACCTGGGTGCGCGCGGTACTACTTTCGTGCAAGAGTGTGCGCTTACGCCACGGT 408
QY 378 GCTGCACATACCGCGCGTACGCGTCGAGCGCTACCTGGCCATCTGCGCGCGCGCTCGCGC 437

```

```

DB 409 GCTGAGCGTGGCAGGCGCTGAGCGCGCGCTAGCGCGTGTGCGAGCCCTGCGTGC 468
QY 438 CGCGGCTTTGGTCAACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTGGGCGCTGGC 497
DB 469 CCGCAGCCTCTGACACCGCACCGCGGACCGCGTGGTGTGGCGCTCTGTTGGGCGCGCTC 528
QY 498 GCTGCTCTCTGCGCGGTCCCTTCTTGTTCCTGTTGGGCGTGCAGCAGGACCCCG 550
DB 529 GCTCGGCGCTCGCCATGCGCCCATGCGCGTCTCATGCGGCGAGACGACGAACCTCG 581

RESULT 11
BM808829
LOCUS       BM808829
DEFINITION AGENCOURT_6617918 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734336
            5', mRNA sequence.
ACCESSION   BM808829
VERSION     BM808829.1  GI:19125652
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1055)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12739 row: h column: 17
            High quality sequence stop: 589.

FEATURES             source
    1..1055
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5734336"
        /tissue_type="hippocampus"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_124"
        /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
        (destroyed); Site_2: NotI; RNA source male hippocampus,
        age 27. Library is oligo-dT primed and directionally
        cloned (EcoRV site is destroyed upon cloning). Average
        insert size 1.4 kb, insert size range 0.9-4 kb. Library is
        normalized and enriched for full-length clones and was
        constructed by C. Gruber (Invitrogen). Research Genetics
        tracking code 012."

BASE COUNT      160 a      376 c      336 g      183 t
ORIGIN
Query Match      9.0%; Score 125.4; DB 12; Length 1055;
Best Local Similarity 55.0%; Pred. No. 7.3e-16;
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGGCGCGGAGCGCGCGTGGCCCGCGCTGCCGCTTGCAGCAGGCGCG 86
DB 53 CAGCAGCCCGCGCGCCCGCGCGCCAGCTCCAAACCCGGGCTGAGCTGACGCCCGGCT 112
QY 87 CTGCTGCGCTTTCCCTCGGGGCGCTGTGTGCGGTGACCGCTGTGCTGTGCTGTT 146
DB 113 GGGCGTGGACACTCGCTCTGCGCGCAAGTGTCTTTCACCGCGCTCTACGCACTCATCTG 172
QY 147 CGTCTGCGGGTGTAGCGGCAACGTG-----GTGACCGTGATGCTGATCGGGCGCTACCG 200

```



```

Db      547 GCTCGGCTCGCCCTGCCCATGGCCGTCAATCATGGGGCAGA 588

RESULT 15
CNS036WH/c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone
216F22 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION
AL230570.1 GI:7889565
VERSION
GSS: genome survey sequence.
KEYWORDS
SOURCE
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE
1 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
  Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
  Saurin,W. and Weissenbach,J.
  Estimate of human gene number provided by genome-wide analysis
  using Tetraodon nigroviridis DNA sequence
  Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645
REFERENCE
2
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
  Fizames,C., Bernot,A., Wincker,P., Brottier,P., Quetier,F.,
  Saurin,W. and Weissenbach,J.
TITLE
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL
MEDLINE
PUBMED
REFERENCE
3 (bases 1 to 941)
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
  Fizames,C., Bernot,A., Wincker,P., Brottier,P., Quetier,F.,
  Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10899143
REFERENCE
3 (bases 1 to 941)
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
  Fizames,C., Bernot,A., Wincker,P., Brottier,P., Quetier,F.,
  Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..941
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="216F22"
/clone_lib="G"
/notes="Genoscope sequence ID : CNS036WH/c-T7"
BASE COUNT
197 a 277 c 327 g 132 t 8 others
ORIGIN

Query Match 8.8%; Score 122.8; DB 29; Length 941;
Best Local Similarity 53.4%; Pred. No. 2.5e-15;
Matches 318; Conservative 2; Mismatches 270; Indels 5; Gaps 3;

Qy      24 CGACGGCCCCGGGGGGCGCGGAGCGCGCGTGGCCGCGCTTGGCGACGAGCG 83
Db      605 CGAGGCCACCGCGGGGACCTGGACGCGCTGTGACGAGGCGGTACTTGGCCAG 546

Qy      84 CCGTGTGTGGCCCTTTCCCTGGGGGGCGCTGGTCCGGTGACCGCTGTGTCCTGCT 143
Db      545 GTACCGGGGGCCCCCGGAGTCCCGCTTCTGCTGGTGTGCTGCTGCTGCTGCTGCT 486

Qy      144 GTTCGTGTGGGGGTGAGCGGACACGCTGATGCTGATCGGGGGCGCTACCGGGA 203
Db      485 CTTCTGTGGGGGTCTGGGGCAACTCCCTGACCTGGCGCGCTCATCTCGCTTCGGGC 426

```

Search completed: January 1, 2004, 02:52:42
 Job time : 3261.52 secs

```

Qy      204 CATGCGGACCAACCAACCACTTTGATCTGGGACGATGGCGGTTCGACCTAC---TCAT 260
Db      425 GAGCGCGACGCCCAACCACTTACTACTCTGTGA-CCTGGCGGCGTCCGACCTGCTGGTGC 367

Qy      261 CTTGCTCGGGCTGCGGTTGGACCTGTACGCGCTCTGGCGCTCGGGCCCTGGGTGTTCCGG 320
Db      366 CTTGCTGGGGATGCCGCTGGAGCTGTACGACCTGTGGCGGAACCTACCCCTTCTGCTGGG 307

Qy      321 GCCGCTGCTTGGCGGCTGTCCCTCTACGTTGGCGGAGGGCTGCACCTACGCCACGCTGCT 380
Db      306 GGAGGGGGGCTGCTACTTCCGGACCTTCTCTGTTTCGAG-ACGGTGTSTTTCGCTCCATCCT 248

Qy      381 GCACATGACCGCGCTCAGCGTACGCGCTACCTGGGCACTGCGCGCCGCTCCGCGCCG 440
Db      247 CAACGTCACGCGCTCAGCGTGGAGCGCTACCTGGCGGTGGTGCACCCGCTCAAGGTCAA 188

Qy      441 CGTCTTGGTCAACCCGCGCGCTCGCGCGCTCATCGCTGTGCTCTGGGCGCGTGGCGCT 500
Db      187 GCACATGGCACTCGCGCGCACGTCAAGAGGTGGTCTCTCTGCTGTGGTGTGCTGCTCAT 128

Qy      501 GCTCTCTGCGCGTCCCTTCTTCTTGTGGTGGCGTGGAGACGCCGCGCATCTCCGT 560
Db      127 GCTGTGCGCGTGGCGGCAACACACGAGCGTGCACGCGATCGAGGSGTGGCGCCCAAGTTCGGC 68

Qy      561 AGTCCCGGGCTCAATGGCACCCGCGGATGCGCTCTCTCGCTCTCTCGCTCTGCTG 615
Db      67 CGGGAGTTTCTCTGCTCGCGCATCTGCTAGGTCCCTCGAAGCTCTGTGCTGCTGCTCG 13

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 48.8546 Seconds
(without alignments)
1254.097 Million cell updates/sec

Title: US-09-719-485-5

Perfect score: 2043

Sequence: 1 MGSPWNGSDPGEGAREPPWP.....WQNLHKGHFADDVLLSVL 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2043	100.0	386	21	AAV54146
2	2043	100.0	386	22	AAV54146
3	1581	77.4	412	21	AAV54145
4	1581	77.4	412	22	AAV54145
5	1581	77.4	412	22	AAV54145
6	1581	77.4	412	23	ABG30936
7	1581	77.4	412	24	ABP81993
8	1581	77.4	501	22	AAV54145
9	1575	77.1	412	21	AAV54146

10	1573	77.0	412	23	ABB09535	Human motilin rece
11	1307	64.0	400	22	AAV54146	Amino acid sequenc
12	1149.5	56.3	271	22	AAV54146	The puffer fish mo
13	673.5	33.0	363	22	AAV54146	Amino acid sequenc
14	673.5	33.0	363	22	AAV54146	Pig growth hormone
15	649	31.8	289	18	AAV19609	Human growth hormo
16	649	31.8	289	18	AAV19611	Human growth hormo
17	649	31.8	289	18	AAV19612	Human growth hormo
18	649	31.8	289	24	ABP81828	Human growth hormo
19	644	31.5	289	18	AAV19218	A canine growth ho
20	644	31.5	349	21	AAV69293	Human G protein-co
21	643	31.5	366	21	AAV90632	Rat growth hormone
22	643	31.5	366	22	AAV70345	Human G-protein co
23	643	31.5	366	22	AAV97376	Human ghrelin rece
24	643	31.5	366	23	ABB09534	A mouse growth hor
25	643	31.5	366	23	ABB09534	Rat growth hormone
26	641	31.4	364	21	AAV54565	Pig growth hormone
27	641	31.4	364	22	AAV97377	Human growth hormo
28	639.5	31.3	353	18	AAV19608	Human growth hormo
29	639.5	31.3	353	18	AAV19612	Human growth hormo
30	639.5	31.3	361	18	AAV19217	Human growth hormo
31	639.5	31.3	362	18	AAV19610	Rat growth hormone
32	637	31.2	364	18	AAV19613	Human mutant G pro
33	637	31.2	366	21	AAV90866	Rat growth hormone
34	625	30.6	364	18	AAV19220	Human growth hormo
35	454	22.2	271	18	AAV19612	Human growth hormo
36	454	22.2	271	18	AAV19219	Human novel cytol
37	355.5	17.4	445	22	AAU68523	Human novel cytol
38	355.5	17.4	445	22	AAU68566	Human G protein-co
39	354.5	17.4	403	21	AAV90838	Human growth hormo
40	354.5	17.4	403	22	AAV44642	Human FM-3. Homo
41	354.5	17.4	403	22	AAV99185	Human GPCR GPR66
42	354.5	17.4	403	23	ABG95158	Human G protein-co
43	354.5	17.4	403	24	ABP81682	Human G-protein co
44	354.5	17.4	426	22	AAE03628	Human mutant G pro
45	352.5	17.3	403	21	AAV90873	

ALIGNMENTS

RESULT 1
AAV54146
ID AAV54146 standard; Protein; 386 AA.

XX AC AAV54146;

XX DT 27-MAR-2000 (first entry)

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1B.

XX DE Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;

XX DE spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;

XX DE functional defect; neurological disorder; scleroderma; colonscopy;

XX DE paraneoplastic syndrome; radiation induced dysmotility; diabetes;

XX DE infection; stress-related motility disorder; psychogenic disorder;

XX DE gastroparesis; gastro-oesophageal reflux disease; constipation;

XX DE chronic idiopathic pseudo obstruction; acute faecal impaction;

XX DE postoperative ileus; gallstones; infantile colic; diarrhoea;

XX DE irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;

XX DE endoscopy; duodenal intubation.

XX OS Homo sapiens.

XX PN WO9964436-A1.

XX PD 16-DEC-1999.

XX PF 08-JUN-1999; 99WO-US12773.

XX PR 12-JUN-1998; 98US-0089098.

XX PA (MERI) MERCK & CO INC.

XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeill D, Howard AD;
 PI Pong S, Smith RG;
 XX WPI; 2000-105868/09.
 DR N-PSDB; AAZ45404.
 XX Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions -
 XX Claim 5; Fig 5; 4pp; English.
 XX The present sequence represents splice variant MTL-R1B of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility
 CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.
 XX Sequence 386 AA;
 SQ Query Match 100.0%; Score 2043; DB 21; Length 386;
 Best Local Similarity 100.0%; Pred. No. 2.9e-172;
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPLGALVPVTAVCLFLVVGSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPLGALVPVTAVCLFLVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPPGLCRSLVYVGE 120
 DB 61 MLIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPPGLCRSLVYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 DB 121 CTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 QY 241 QLGLARVLMVWTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAAASGRGRHQTIVRL 300
 DB 241 QLGLARVLMVWTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAAASGRGRHQTIVRL 300
 QY 301 RKWSRRGSKDACLOSPGTAQTIGPLPLAQIWAQPLPAPFPISIPASTRRGGSGIYNL 360
 DB 301 RKWSRRGSKDACLOSPGTAQTIGPLPLAQIWAQPLPAPFPISIPASTRRGGSGIYNL 360
 QY 361 LVALPRQNHHLKHGREFADDVLLSVL 386
 DB 361 LVALPRQNHHLKHGREFADDVLLSVL 386
 RESULT 2
 AAB62653
 ID AAB62653 standard; Protein; 386 AA.
 XX AAB62653;
 AC AAB62653;
 XX 23-JUL-2001 (first entry)
 DT Short form of motilin receptor, GPR-38B isoform.
 XX

XX zsig33; signal transduction; hormone; enzyme; neural development;
 KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 KW glucose; osteopathic; anorectic; vulnary; immunomodulator; GHS-R;
 KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.
 XX Homo sapiens.
 OS WO200138355-A2.
 PN 31-MAY-2001.
 PD 22-NOV-2000; 2000WO-US32074.
 XX 22-NOV-1999; 99US-0166765.
 PR (ZYMO) ZYMOGENETICS INC.
 PA Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
 PI WPI; 2001-355879/37.
 DR N-PSDB; AAF83684.
 XX Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide -
 XX Disclosure; Page 106-109; ilpp; English.
 PS The invention relates to a method of forming a reversible peptide-
 XX receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
 CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the short form of
 CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result
 CC from alternative splicing). GPR38 has homology to the human G-protein
 CC coupled receptor, GHS-R.
 XX Sequence 386 AA;
 SQ Query Match 100.0%; Score 2043; DB 22; Length 386;
 Best Local Similarity 100.0%; Pred. No. 2.9e-172;
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPLGALVPVTAVCLFLVVGSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPLGALVPVTAVCLFLVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPPGLCRSLVYVGE 120
 DB 61 MLIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPPGLCRSLVYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 DB 121 CTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240

Db 181 QDGISSVPLGNGTARASSPLASSPPLWLSRAPPPSPGPGTAAALFSSRECRPSA 240
Qy 241 QLGALRVMLVWTTAYFFLPCLCLSLYLGLIGRELWSSRRRLGPAASGRGRHQRQTVRVL 300
Db 241 QLGALRVMLVWTTAYFFLPCLCLSLYLGLIGRELWSSRRRLGPAASGRGRHQRQTVRVL 300
Qy 301 RKNRSRSGKDACIQSAPPGTAQTGLPLLAQWAPLPAPFPPISSIPASTRRGGSGIYNL 360
Db 301 RKNRSRSGKDACIQSAPPGTAQTGLPLLAQWAPLPAPFPPISSIPASTRRGGSGIYNL 360
Qy 361 LVALPRQNLHKGREADDVLLSVL 386
Db 361 LVALPRQNLHKGREADDVLLSVL 386

RESULT 3
AAY54145
ID AAY54145 standard; Protein; 412 AA.
XX AC AAY54145;
XX DT 27-MAR-2000 (first entry)
XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.
XX KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
KW endoscopy; duodenal intubation.
XX OS Homo sapiens.
XX PN WO9964436-A1.
XX PD 16-DEC-1999.
XX PF 08-JUN-1999; 99WO-US12773.
XX PR 12-JUN-1998; 98US-0089098.
XX PA (MERI) MERCK & CO INC.
XX PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI Pong S, Smith RG;
XX DR WPI; 2000-105868/09.
XX DR N-PSDB; AAZ45403.
XX PT Novel receptor protein for screening compounds used in treating
PT irritable bowel syndrome, constipation and other gastric conditions
XX PS Claim 3; Fig 3; 44pp; English.
XX CC The present sequence represents splice variant MTL-R1A of the motilin
CC receptor. The gene encodes a G-protein coupled receptor, and is
CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
CC functional seven transmembrane domain form, and MTL-R1B is a truncated
CC five transmembrane domain. The MTL-R1 proteins are used to identify
CC agonists and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathic pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile colic, irritable

CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX Sequence 412 AA;
Qy 1 MGSPWNGSDGPGAREPPMPALPPCDERRCSPPPLGALVPTAVCLCLFVGVSGNNVTV 60
Db 1 MGSPWNGSDGPGAREPPMPALPPCDERRCSPPPLGALVPTAVCLCLFVGVSGNNVTV 60
Qy 61 MLIGRYRDMRTTNNLYLGSMAVSDLLILGLPFDLYLRWSPWVFGPCLLRLSLVVGEG 120
Db 61 MLIGRYRDMRTTNNLYLGSMAVSDLLILGLPFDLYLRWSPWVFGPCLLRLSLVVGEG 120
Qy 121 CTYATLLHMTALSVRYLAICRPLRARVLTTRRRVRALIAVLWAVALLSAGPPLFLVGVE 180
Db 121 CTYATLLHMTALSVRYLAICRPLRARVLTTRRRVRALIAVLWAVALLSAGPPLFLVGVE 180
Qy 181 QDGISSVPLGNGTARASSPLASSPPLWLSRAPPPSPGPGTAAALFSSRECRPSA 240
Db 181 QDGISSVPLGNGTARASSPLASSPPLWLSRAPPPSPGPGTAAALFSSRECRPSA 240
Qy 241 QLGALRVMLVWTTAYFFLPCLCLSLYLGLIGRELWSSRRRLGPAASGRGRHQRQTVRVL 300
Db 241 QLGALRVMLVWTTAYFFLPCLCLSLYLGLIGRELWSSRRRLGPAASGRGRHQRQTVRVL 300

RESULT 4
AAB62652
ID AAB62652 standard; Protein; 412 AA.
XX AC AAB62652;
XX DT 23-JUL-2001 (first entry)
XX DE Long form of motilin receptor, GPR-38A isoform.
XX KW zsig33; signal transduction; hormone; enzyme; neural development;
KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
KW Glucose; osteopathic; anorectic; vulnerary; immunomodulator; GHS-R;
KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.
XX OS Homo sapiens.
XX PN WO200138355-A2.
XX PD 31-MAY-2001.
XX PF 22-NOV-2000; 2000WO-US32074.
XX PR 22-NOV-1999; 99US-0166765.
XX PA (ZYMO) ZYMOGENETICS INC.
XX PI Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
XX DR WPI; 2001-355879/37.
XX DR N-PSDB; AAF83683.
XX PT Forming reversible peptide receptor complex for purifying cell and
PT peptides, stimulating signal transduction and modulating hormone
PT secretion, involves contacting a receptor with zsig33 polypeptide
XX PS Disclosure; Page 102-104; 111pp; English.
XX CC The invention relates to a method of forming a reversible peptide-
CC receptor complex that involves providing an immobilized receptor, and
CC contacting the receptor with a zsig33 peptide (comprising residues 24-37

of AAB62649), where the receptor binds to the zsig33 peptide. The method is useful for purifying cells, purifying a peptide, stimulating signal transduction in a cell expressing a receptor. It is also useful for modulating secretion of hormones, neural development and/or utilization, gastric contractility, nutrient uptake, secretion of digestive and pancreatic enzymes and hormones, secretion of insulin-like growth factor -1, secretion of non-zsig33 proteins. It is useful for modulating growth hormone secretion in a mammal having a disease associated with abnormal levels of growth hormone, such as osteoporosis, bone repair, bone remodeling, low osteoblast levels, cartilage repair and remodeling, skeletal dysplasia, immune suppression, obesity, growth retardation, protein catabolic responses after surgery, cachexia, protein loss, dwarfism, wound healing and ovulation induction, treating a mammal having a metabolic disorder requiring neurological feedback, such as satiety regulation, glucose absorption and metabolism and neuropathy-associated gastrointestinal disorders, and stimulating glucose-induced insulin release in a mammal. The present sequence represents the long form of motilin receptor, GPR38A (one of the two isoforms of GPR38 which result from alternative splicing). GPR38 has homology to the human G-protein coupled receptor, GHS-R.

XX SQ Sequence 412 AA;
Query Match 77.4%; Score 1581; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGEGAREPPWPPALPCDERRCSPPLGALVPVTA VCLCLFVVGSGNVTV 60
DB 1 MGSPWNGSDGEGAREPPWPPALPCDERRCSPPLGALVPVTA VCLCLFVVGSGNVTV 60
QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPDLRLWRSRPWFGLCLRLSLYVGE 120
DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPDLRLWRSRPWFGLCLRLSLYVGE 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
DB 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
QY 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240
DB 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240
QY 241 QLGAIRVLMVTTAYFFLPFLCLSLYGLIGRLWSSRRPLRGPAAASGRGRHQTVRVL 300
DB 241 QLGAIRVLMVTTAYFFLPFLCLSLYGLIGRLWSSRRPLRGPAAASGRGRHQTVRVL 300

RESULT 5

AAB68478
ID AAB68478 standard; Protein; 412 AA.

XX AAB68478;

AC 23-JUL-2001 (first entry)

DT Amino acid sequence of a human motilin receptor polypeptide.

DE Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea.

XX Homo sapiens.

OS WO200132710-A1.

PN 10-MAY-2001.

XX 25-OCT-2000; 2000WO-US29426.

XX 29-OCT-1999; 99US-0162264.

XX (MERI) MERCK & CO INC.

XX

PI Tan C, McKee K;
XX WPI; 2001-343479/36.
DR N-PSDB; AAF85449.
DR Novel polypeptides related to dog and rabbit motilin receptor
XX polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX Disclosure; Page 32-33; 42pp; English.

XX The present sequence represents a human motilin receptor polypeptide.
XX The specification describes a unique sequence present in exon 1 of
CC the dog motilin receptor, which is not present in human or Sphaeroides
CC nephelus 7587 motilin receptor sequences. The unique nucleic acid
CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.

XX SQ Sequence 412 AA;
Query Match 77.4%; Score 1581; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGAREPPWPPALPCDERRCSPPLGALVPVTA VCLCLFVVGSGNVTV 60

DB 1 MGSPWNGSDGEGAREPPWPPALPCDERRCSPPLGALVPVTA VCLCLFVVGSGNVTV 60

QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPDLRLWRSRPWFGLCLRLSLYVGE 120

DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPDLRLWRSRPWFGLCLRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180

DB 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180

QY 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240

DB 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240

QY 241 QLGAIRVLMVTTAYFFLPFLCLSLYGLIGRLWSSRRPLRGPAAASGRGRHQTVRVL 300

DB 241 QLGAIRVLMVTTAYFFLPFLCLSLYGLIGRLWSSRRPLRGPAAASGRGRHQTVRVL 300

RESULT 6

ABG30936
ID ABG30936 standard; Protein; 412 AA.

XX ABG30936;

XX 21-OCT-2002 (first entry)

DT Human G protein-coupled receptor 38 (GPR38).

DE Human; G protein-coupled receptor 38; receptor; GPR38;

XX Alzheimer's disease; Parkinson's disease; ulcerative colitis;

XX Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;

XX colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;

XX pancreatic small cell carcinoma; pancreatic adenocarcinoma.

OS Homo sapiens.

XX Key Location/Qualifiers

PH Region 15..31

FT /note= "Antigenic fragment"

FT 217..232

FT /note= "Antigenic fragment"

CC	Best Local Similarity 100.0%; Pred. No. 2.2e-131;	CC	produce antibodies immunospecific for the polypeptides, and to identify
XX	Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	membrane bound or soluble receptors.
SQ	Sequence 501 AA;	SQ	Sequence 501 AA;
QY	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTA	QY	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTA
Db	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTA	Db	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTA
QY	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWF	QY	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWF
Db	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWF	Db	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWF
QY	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVL	QY	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVL
Db	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVL	Db	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVL
QY	181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPG	QY	181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPG
Db	181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPG	Db	181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPG
QY	241 QLGALRVMLVMTTAYFFLPFLCLISILYGLIGRELWSSRR	QY	241 QLGALRVMLVMTTAYFFLPFLCLISILYGLIGRELWSSRR
Db	241 QLGALRVMLVMTTAYFFLPFLCLISILYGLIGRELWSSRR	Db	241 QLGALRVMLVMTTAYFFLPFLCLISILYGLIGRELWSSRR
RESULT 8		RESULT 9	
AAG65822		AAB02854	
ID	AAG65822 standard; Protein; 501 AA.	ID	AAB02854 standard; Protein; 412 AA.
XX	AC AAG65822;	XX	AC AAB02854;
XX	30-JAN-2002 (first entry)	XX	22-AUG-2000 (first entry)
DE	Human GPR38 variant GPR38V polypeptide.	DE	Human G protein coupled receptor hGPR38 (V297K) protein SEQ ID NO:130.
KW	GPR38V; variant; antibacterial; cytostatic; analgesic; antiashtmatic;	KW	Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW	anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;	KW	identification; agonist; screening; therapeutic; pharmaceutical;
KW	antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;	KW	mutant.
KW	antiulcer; antiemetic; cardiant; vaccine; human.	OS	Homo sapiens.
OS	Homo sapiens.	OS	Synthetic.
PN	WO200164836-A2.	PN	WO200022131-A2.
XX	07-SEP-2001.	XX	20-APR-2000.
XX	28-FEB-2001; 2001WO-US06277.	XX	13-OCT-1999; 99WO-US24065.
XX	01-MAR-2000; 2000US-0516315.	XX	13-OCT-1998; 98US-0170496.
PA	(SMIK) SMITHKLINE BEECHAM CORP.	XX	12-NOV-1998; 98US-0108029.
PI	Elshourbagy N, Shabon U;	XX	20-NOV-1998; 98US-0109213.
XX	WPI; 2001-638956/73.	XX	27-NOV-1998; 98US-0110060.
XX	N-PSDB; AAI66989.	XX	16-FEB-1999; 99US-0120416.
PT	New human GPR38V polypeptide and polynucleotide, useful for treating	XX	26-FEB-1999; 99US-0121852.
PT	e.g. bacterial, fungal, protozoal and viral infections, cancers or	XX	12-MAR-1999; 99US-0123944.
PT	allergies, as vaccines, and for identifying agonists and antagonists	XX	12-MAR-1999; 99US-0123945.
XX	potentially useful in therapy	XX	12-MAR-1999; 99US-0123946.
PS	Claim 1; Page 26; 32pp; English.	XX	12-MAR-1999; 99US-0123948.
XX		XX	12-MAR-1999; 99US-0123949.
CC	This represents a human GPR38 variant (GPR38V) polypeptide. GPR38V can	XX	12-MAR-1999; 99US-0123951.
CC	be expressed by standard recombinant methodology. The polynucleotides and	XX	28-MAY-1999; 99US-0136436.
CC	polypeptides are used in the treatment of bacterial, fungal, protozoal	XX	28-MAY-1999; 99US-0136437.
CC	and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,	XX	28-MAY-1999; 99US-0136439.
CC	diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart	XX	28-MAY-1999; 99US-0137127.
CC	failure, hypertension, urinary retentions, osteoporosis, allergies,	XX	28-MAY-1999; 99US-0137131.
CC	ulcers, migraine, psychotic and neurological disorders, or dyskinesias.	XX	28-MAY-1999; 99US-0137567.
CC	They are also useful for identifying agonists and antagonists that are	XX	30-JUN-1999; 99US-0141448.
CC	potentially useful in therapy, as vaccines to induce immunological	XX	27-AUG-1999; 99US-0151114.
CC	response in a mammal. The polypeptides may also be used as immunogens to		

PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156634.
 PA (AREN-) ARENA PHARM INC.
 XX Behan DP, Lehmann-Bruinema K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX WPI: 2000-317986/27.
 DR N-PSDB; AAA46116.
 XX Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX Example 2; Page 168-169; 187pp; English.
 XX The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (Orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX Sequence 412 AA;
 SQ
 Query Match 77.1%; Score 1575; DB 21; Length 412;
 Best Local Similarity 99.7%; Pred. No. 7.4e-131;
 Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFLGALVPVTVAVCLCLFVGVSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFLGALVPVTVAVCLCLFVGVSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPFLCLSLYVGE 120
 DB 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPFLCLSLYVGE 120
 QY 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLWALLSAGPFLVGE 180
 DB 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLWALLSAGPFLVGE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 QY 241 QLGAALRVMLWTVTAYFFLPFLCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 DB 241 QLGAALRVMLWTVTAYFFLPFLCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 RESULT 10
 ABB09535 standard; protein; 412 AA.
 XX ABB09535;
 AC ABB09535;
 XX 22-OCT-2002 (first entry)
 DE Human motilin receptor.
 XX Human; motilin receptor; appetite; food intake; agonist; analogue;
 KW undernutrition; anorexia; cachexia; malignant disease; infection;
 KW inflammatory disease; weight loss; antagonist; obesity; anorectic;
 KW anabolic; ghrelin receptor homologue; receptor.
 OS Homo sapiens.
 XX WO200260472-A1.

XX 08-AUG-2002.
 PD 31-JAN-2002; 2002WO-JP00765.
 PF 31-JAN-2001; 2001JP-0024423.
 PR (CHUS) CHUGAI SEIYAKU KK.
 PA Inui A, Asakawa A, Kaga T;
 PI WPI: 2002-619206/66.
 XX Remedies for diseases with hypo-nutrition status e.g. inappetence and
 PT cachexia, containing ghrelin or its analogs including agonists and
 PT antagonists -
 XX Disclosure; Fig 1B; 50pp; Japanese.
 XX The invention relates to the use of ghrelin or its analogues for the
 CC treatment of diseases associated with undernutrition such as anorexia,
 CC and also relates to the use of ghrelin antagonists for the prevention
 CC or treatment of obesity. The invention additionally discloses a method
 CC for screening ghrelin agonists or antagonists by measuring the amount
 CC of food intake, neuropeptide Y (NPY) expression, binding of NPY to
 CC NPY receptor Y1, oxygen consumption, gastric emptying, or activity of
 CC the vagus nerve. Intracerebroventricular (ICV) administration of ghrelin
 CC in animals was found to increase food intake over a period of 24 hours.
 CC Ghrelin and its analogues may therefore be used to treat conditions
 CC such as loss of appetite, anorexia, cachexia, malignant diseases, and
 CC weight loss associated with infection or inflammatory diseases.
 CC Conversely, ghrelin antagonists may be used in the treatment of obesity.
 CC The present sequence represents the human motilin receptor, a homologue
 CC of the ghrelin receptor (ABB09534) which is referred to in the disclosure
 CC of the invention.
 XX Sequence 412 AA;
 SQ
 Query Match 77.0%; Score 1573; DB 23; Length 412;
 Best Local Similarity 99.7%; Pred. No. 1.1e-130;
 Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFLGALVPVTVAVCLCLFVGVSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFLGALVPVTVAVCLCLFVGVSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPFLCLSLYVGE 120
 DB 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPFLCLSLYVGE 120
 QY 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLWALLSAGPFLVGE 180
 DB 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLWALLSAGPFLVGE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 QY 241 QLGAALRVMLWTVTAYFFLPFLCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 DB 241 QLGAALRVMLWTVTAYFFLPFLCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 RESULT 11
 AAB68477
 ID AAB68477 standard; Protein; 400 AA.
 XX AAB68477;
 AC AAB68477;
 XX 23-JUL-2001 (first entry)
 DE Amino acid sequence of a rabbit motilin receptor polypeptide.
 XX

KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 OS Oryctolagus cuniculus.
 XX WO200132710-A1.
 PN 10-MAY-2001.
 PD 25-OCT-2000; 2000WO-US29426.
 PF 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 PA Tan C, McKee K;
 PI WPI; 2001-343479/36.
 XX DR N-PSDB; AAF85448.
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX Claim 8; Page 18; 42pp; English.
 PS The present sequence represents a rabbit motilin receptor polypeptide.
 CC The specification describes a unique sequence present in exon 1 of
 CC the dog motilin receptor, which is not present in human or Sphaeroides
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX Sequence 400 AA;

Query Match 64.0%; Score 1307; DB 22; Length 400;
 Best Local Similarity 85.3%; Pred. No. 3.5e-107;
 Matches 256; Conservative 5; Mismatches 27; Indels 12; Gaps 3;
 QY 1 MGSPWNGSGPGAREPPWPPALPPCDERRCSPFFLGALVPVTAVCICLPVVGSGNVTV 60
 Db 1 MGSPWNGSGPGAREPPWPPALPPCDERRCSPFFLGALVPVTAVCICLPVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYLRWSPVWFGPPLCRSLYVGE 120
 Db 61 LLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYLRWSPVWFGPPLCRSLYVGE 120
 QY 121 CTYATLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 Db 121 CTYASLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 QY 181 QDPGISVVGNGTARIASSPLASPPWLSRAPPPSPGPEAAALFSCRCRPSA 240
 Db 181 QDPAPVAPDRNGTV-----PLDPSG-----APASPPSGP-GAAALFSCRCRPSA 228
 QY 241 QLGLRVMLWVTYFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 Db 229 QLGLRVMLWVTYFFFLPFLCLSLYGLIAQLWRGRLGPAATGRGRHQTVRVL 288

RESULT 12
 AAB68476
 ID AAB68476 standard; Protein; 271 AA.
 XX AAB68476;
 AC AAB68476;
 XX 23-JUL-2001 (first entry)
 DT 23-JUL-2001 (first entry)
 XX

DE Amino acid sequence of a dog motilin receptor exon 1.
 XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 XX Canis sp.
 XX WO200132710-A1.
 PN 10-MAY-2001.
 PD 25-OCT-2000; 2000WO-US29426.
 PF 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 PA Tan C, McKee K;
 PI WPI; 2001-343479/36.
 XX DR N-PSDB; AAF85447.
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX Claim 4; Page 17; 42pp; English.
 PS The present sequence is encoded by exon 1 of a dog motilin receptor gene.
 CC The specification describes a unique sequence present in exon 1 of
 CC the motilin receptor, which is not present in human or Sphaeroides
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX Sequence 271 AA;

Query Match 56.3%; Score 1149.5; DB 22; Length 271;
 Best Local Similarity 77.0%; Pred. No. 1.9e-93;
 Matches 231; Conservative 13; Mismatches 27; Indels 29; Gaps 4;
 QY 1 MGSPWNGSGPGAREPPWPPALPPCDERRCSPFFLGALVPVTAVCICLPVVGSGNVTV 60
 Db 1 MGSPWNGSGPGAREPPWPPALPPCDERRCSPFFLGALVPVTAVCICLPVVGSGNVTV 53
 QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYLRWSPVWFGPPLCRSLYVGE 120
 Db 54 LLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYLRWSPVWFGPPLCRSLYVGE 113
 QY 121 CTYATLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLVGE 180
 Db 114 CTYATLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLVGE 173
 QY 181 QDPGISVVGNGTARIASSPLASPPWLSRAPPPSPGPEAAALFSCRCRPSA 240
 Db 174 QDAG---GFLGNGSARLARA-----PSPPPGPE-----AALFSCRCRPSA 211
 QY 241 QLGLRVMLWVTYFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 Db 212 QLGLRVMLWVTYFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 271

RESULT 13
 AAY54147
 ID AAY54147 standard; Protein; 363 AA.
 XX AAY54147;
 AC AAY54147;
 XX

[illegible]

RESULT 2

US-09-077-674-10
 ; Sequence 10, Application US/09077674
 ; Patent No. 6531314
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feighner, Scott D.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Liberator, Paul A.
 ; APPLICANT: Schaeffer, James M.
 ; APPLICANT: Van Der Ploeg, Leonardus
 ; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/077,674
 ; FILING DATE: 3-JUN-1998
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cocuzzo, Anna L.
 ; REGISTRATION NUMBER: 42,452
 ; REFERENCE/DOCKET NUMBER: 19589P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1273
 ; TELEFAX: 732-594-4720
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 289 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single

[illegible]

RESULT 3

US-09-077-675A-13
Sequence 13, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
INFORMATION FOR SEO ID NO: 13:

SEQUENCE CHARACTERISTICS:
 LENGTH: 366 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-077-675A-13

Query Match 31.5%; Score 643; DB 3; Length 366;
 Best Local Similarity 45.1%; Pred. No. 1.1e-45;
 Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGVSNG 56
 DB 2 WNAATPSEEPGFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALLFVVGIA 61
 QY 57 VVTVMILIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWVFGPPLLCRLSLY 116
 DB 62 LITMLVVSFRERLTTNNLYLSSMAFSDLLIFLCMLDLVRLWQYRPWFGDLLCKLQF 121
 QY 117 VBGCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIALVLAVALLSAGPFLFL 176
 DB 122 VSESCYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
 QY 177 VGVQDPGIVVPELNGTARIASSPPLWLSRAPPPSPSPGPETAABAAALFRRER 236
 DB 182 VGVHE-----NGT-----DP--W-----DTNECR 199
 QY 237 PS--PAQLGALRYMLVTTAYFPFLPCLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294
 DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCULTVLSLIGRKLWRRRRGDVVGASLRDQNHK 259
 QY 295 QTVRVL 300
 DB 260 QTVKML 265

RESULT 4

US-09-077-674-13
 Sequence 13, Application US/09077674
 Patent No. 6531314

GENERAL INFORMATION:

APPLICANT: Arena, Joseph P.
 APPLICANT: Cully, Doris F.
 APPLICANT: Feighner, Scott D.
 APPLICANT: Howard, Andrew D.
 APPLICANT: Liberator, Paul A.
 APPLICANT: Schaeffer, James M.
 APPLICANT: Van Der Ploeg, Leonardus
 TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merck & Co., Inc.
 STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 CITY: Rahway
 STATE: NJ
 COUNTRY: USA
 ZIP: 07065-0900
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/077, 674
 FILING DATE: 3-JUN-1998
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Cocuzzo, Anna L.
 REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19589P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 732-594-1273
 TELEFAX: 732-594-4720
 TELEX:
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 366 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-077-674-13

Query Match 31.5%; Score 643; DB 4; Length 366;
 Best Local Similarity 45.1%; Pred. No. 1.1e-45;
 Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGVSNG 56
 DB 2 WNAATPSEEPGFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALLFVVGIA 61
 QY 57 VVTVMILIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWVFGPPLLCRLSLY 116
 DB 62 LITMLVVSFRERLTTNNLYLSSMAFSDLLIFLCMLDLVRLWQYRPWFGDLLCKLQF 121
 QY 117 VBGCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIALVLAVALLSAGPFLFL 176
 DB 122 VSESCYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
 QY 177 VGVQDPGIVVPELNGTARIASSPPLWLSRAPPPSPSPGPETAABAAALFRRER 236
 DB 182 VGVHE-----NGT-----DP--W-----DTNECR 199
 QY 237 PS--PAQLGALRYMLVTTAYFPFLPCLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294
 DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCULTVLSLIGRKLWRRRRGDVVGASLRDQNHK 259
 QY 295 QTVRVL 300
 DB 260 QTVKML 265

RESULT 5

US-09-170-496D-88
 Sequence 88, Application US/09170496D
 Patent No. 6555339
 GENERAL INFORMATION:
 APPLICANT: Behan, Dominic P.
 APPLICANT: Chalmers, Derek T.
 APPLICANT: Liaw, Chen W.
 TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
 TITLE OF INVENTION: Receptors
 FILE REFERENCE: AREN-0040
 CURRENT APPLICATION NUMBER: US/09/170,496D
 CURRENT FILING DATE: 1998-10-13
 NUMBER OF SEQ ID NOS: 294
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 88
 LENGTH: 366
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-170-496D-88

Query Match 31.5%; Score 643; DB 4; Length 366;
 Best Local Similarity 45.1%; Pred. No. 1.1e-45;
 Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGVSNG 56
 DB 2 WNAATPSEEPGFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALLFVVGIA 61
 QY 57 VVTVMILIGRYDRMTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWVFGPPLLCRLSLY 116

Db 62 LTLMLVVSRRFRLTNNLYLSSMAFSDLLIFLCMPDLVRLMQYRWNFGDLICKLFQF 121
QY 117 VGEQCTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLALIAVLWAVALLSAGPFLFL 176
Db 122 VSSCTYATVITLITLALSVRYFAICFPLRAKVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLASPPMLSRAPPPSPGPTAEAAALFSRECR 236
Db 182 VGVHEH-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRRLRGPAAAGREGRH 294
Db 200 PTEFAVRSGLLTVMWVSSIFFLPFLVCLTVLYSLIGRKLWRRRRRGDAVVGASLRDQNHK 259
QY 295 QTVRVL 300
Db 260 QTVKML 265

RESULT 6

US-09-077-675A-5
; Sequence 5, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-5

Query Match 31.4%; Score 641; DB 3; Length 289;
Best Local Similarity 43.3%; Pred. No. 1.2e-45;
Matches 142; Conservative 45; Mismatches 77; Indels 64; Gaps 9;
QY 5 WNGSDGPEGA-----REPPWPALPPCD---ERRCSPPPLGALVPVTAICLCFLFWGVSGN 56

Db 2 WNAATPSEEPGNLTLPDLGWDAPPENDSLVEELLPLFPPTLLAGVTATCVALFVVGIAGN 61
QY 57 VVTVMILIGRYDMMRTTNNLYLSSMAVSDLLILGLPFDLYRLWRSRPWVPGPLLCRLSLY 116
Db 62 LTLMLVVSRRFRLTNNLYLSSMAFSELLIFLCMPLEFLRMLQYRWNFGDLICKLFQF 121
QY 117 VGEQCTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLALIAVLWAVALLSAGPFLFL 176
Db 122 VSSCTYATVITLITLALSVRYFAICFPLRAKVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLASPPMLSRAPPPSPGPTAEAAALFSRECR 236
Db 182 VGVHEH-----NGT-----DPRD-----TNECR 199
QY 237 PS--PAQGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRRLRGPAAAG---RER 291
Db 200 ATEFAVRSGLLTVMWVSSIFFLPFLVCLTVLYSLIGRKLW---RRKGEAAVGSRLRDQ 256
QY 292 GHRQTVRLRKWSRRGSKDACLSQAPPG 319
Db 257 NHQTVKML-----GGSQCALELSLPG 278

RESULT 7

US-09-077-674-5
; Sequence 5, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-5

```
Query Match 31.4%; Score 641; DB 4; Length 289;
Best Local Similarity 43.3%; Pred. NO. 1.2e-45;
Molecule Type: protein
Matches 142; Conservative 45; Mismatches 77; Indels 64; Gaps 9;

QY 5 WNGSDGEGA-----REPPALPPCD---ERRCSPPGLGALVPVTAVCCLFVVGSGN 56
DB 2 WNATPSEEPGNUTLPDLGWDAPPENDSLVEELLPLFPFTLLAGVTATCVALFVVGIGN 61
QY 57 VVTVMILIGRYDRMRTTNLYLGSMAVSDDLILGLPDLVRLWRSRPWFGLLCRLSL 116
DB 62 LLTMLVRSRFRMTTNLYLSSMAFSELLIFLCMPLEFLWQYRPNWNLNLLCKLQF 121
QY 117 VGGCTYATLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLWAVALLSAGPLFL 176
DB 122 VSESCTYATLITLALSVERYFAICFPLRAKVVVTKGRVKLVILVIAVAFCSAGPIFVL 181
QY 177 VGVQDQPGISVVPGLNCTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECR 236
DB 182 VGVHED-----NGT-----DPRD-----TNECR 199
QY 237 PS--PAQGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPASG---RER 291
DB 200 ATEFAVRSGLLTVMVWSSVFFFLPVFCVLTVLSLIGRKLW---RRKGEAAVGSLLRQ 256
QY 292 GHRQTVRVLRKWSRSGKACLOSAPFG 319
DB 257 NHKQTVKML-----GGSQCALELSLPG 278

RESULT 8
US-09-719-485-5-16
; Sequence 16, Application US/09077675A
; Patent No. 6243199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ FOR WINDOWS Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-16

Query Match 31.4%; Score 641; DB 3; Length 364;
Best Local Similarity 45.8%; Pred. NO. 1.5e-45;
Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

QY 5 WNGSDGEGA---PWPALPPCD---ERRCSPPFPPLGALVPVTAVCCLFVVGSG 55
DB 2 WNAT--PSEPEPNVTLDLDWDASPCNDLSLPDELLPLFPAPLLAGVTATCVALFVVGISG 59
QY 56 NVVTVMILIGRYDRMRTTNLYLGSMAVSDDLILGLPDLVRLWRSRPWFGLLCRLSL 115
DB 60 NLLTMLVRSRFRMTTNLYLSSMAFSDLLIFLCMPDLVRLWQYRPNWNLNLLCKLQF 119
QY 116 YVGGCTYATLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLWAVALLSAGPLFL 175
DB 120 VSESCTYATLITLALSVERYFAICFPLRAKVVVTKGRVKLVILVIAVAFCSAGPIFV 179
QY 176 LVGVQDQPGISVVPGLNCTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECR 235
DB 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPASG---RE 290
DB 198 RATEFAVRSGLLTVMVWSSVFFFLPVFCVLTVLSLIGRKLW---RR--RGDAAVGASLRD 253
QY 291 RGHRTVTVRL 300
DB 254 QNHKQTVKML 263

RESULT 9
US-09-719-485-5-16
; Sequence 16, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris P.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ FOR WINDOWS Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
```

```

;
; TELEFAX: 732-594-4720
;
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 16:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-077-674-16
;
;
; Query Match 31.4%; Score 641; DB 4; Length 364;
; Best Local Similarity 45.8%; Pred. No. 1.5e-45;
; Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;
;
;
; QY 5 WNGSDQPEGAREP-----PWPALPPCD---ERRCSFPPLGALVPVTVAVCLCLFVVGVS 55
; DB 2 WNAT--PSEPEPNVTLDDWDASPGNDSPDELPLFPAPPLAGVTATCVAFVVGISG 59
;
; QY 56 NVVTVMILIGYRDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSL 115
; DB 60 NLLTMLVRSFRRLRTTNNLYLSSMAPSDLLIFLCMPDLVRLWQYRPNWFGDLLCKLFQ 119
;
; QY 116 YVCEGCTYATLHWTALSVERYLAIQPIARVLVTRRRVRALIAVLWVALISAGPFLF 175
; DB 120 FVSECTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVILVTIWAFAFCAGPIFV 179
;
; QY 176 LVGVEODPGISVVGVLGNTGARIASSPLASSPPLWLSRAPPPSPGSETAAALFSGEC 235
; DB 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
;
; QY 236 RPS--PAQLGALRVMLVMTVTTAYFFLPFLCLISLYGLIGRELWSSRRRLRGPAAAG---RE 290
; DB 198 RATEFAVRSGLLTVMVWVSSVFFFLPVFCLTVLYSLIGRKLW--RR--RGDAAVGASLIRD 253
;
; QY 291 RGHRTQVRVL 300
; DB 254 QNHKQTVKML 263
;
;
; RESULT 10
;
; US-09-077-675A-3
; Sequence 3, Application US/09077675A
; Patent No. 6242199
;
; GENERAL INFORMATION:
;
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.

```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-3

Query Match 31.3%; Score 639.5; DB 4; Length 353;
Best Local Similarity 47.2%; Pred. No. 2e-45;
Matches 137; Conservative 39; Mismatches 61; Indels 53; Gaps 7;

QY 19 WPAIPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGSGNVVTVMILIGRYDRMTITNL 75
Db 8 WDAPPENDSLVEELLPLFPPTPLLAGVTATCVAFVVGAGNLLTMLVVSFRERMTITNL 67
QY 76 YLGSMAVSDLLILGLPFDLYLRWSRPVWFGPGLLCRLSLYVGEGETYATLLHMTALSVE 135
Db 68 YLSSMAFSDLLIFLCPLDLFRLWQVRPNWLNGLLCKLFQVSESCYATVLTITALSVE 127
QY 136 RYLAICRPLARVLTFRVRLAIVLWAVALLSAGPFLVGVQDPGISVVPGLNGTA 195
Db 128 RYFAICFPLRAKVVTGKRVKLIVLWAVAFCSAGPIFVLVGEHDE-----NGT- 177
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS--PAQLGALRVMLWVTT 253
Db 178 -----DPRD-----TNECRTEFAVRSGLLTVMWVSS 205
QY 254 AYFFLPFLCLSLYGLIGRELMSRRPLRGPAAAG---RERGHRTVRL 300
Db 206 VFFFLPVFCLTVLYSLIGRLKLM---RRKRGAAVGSRLRDQNHKQTVKML 252

RESULT 12
US-09-077-675A-8
; Sequence 8, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-8

Query Match 31.3%; Score 639.5; DB 3; Length 361;
Best Local Similarity 47.0%; Pred. No. 2e-45;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

QY 19 WPAIPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGSGNVVTVMILIGRYDRMTITNL 75
Db 16 WDASPGNDSIGDELLQLPAPLLAGVTATCVAFVVGAGNLLTMLVVSFRERLTITNL 75
QY 76 YLGSMAVSDLLILGLPFDLYLRWSRPVWFGPGLLCRLSLYVGEGETYATLLHMTALSVE 135
Db 76 YLSSMAFSDLLIFLCPLDLFRLWQVRPNWLNGLLCKLFQVSESCYATVLTITALSVE 135
QY 136 RYLAICRPLARVLTFRVRLAIVLWAVALLSAGPFLVGVQDPGISVVPGLNGTA 195
Db 136 RYFAICFPLRAKVVTGKRVKLIVLWAVAFCSAGPIFVLVGEHDE-----NGT- 185
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS--PAQLGALRVMLWVTT 253
Db 186 -----DP--W-----DTNECRTEFAVRSGLLTVMWVSS 213
QY 254 AYFFLPFLCLSLYGLIGRELMSRRPLRGPAAAGRERGHRTVRL 300
Db 214 IFFFLPVFCLTVLYSLIGRLKLMRRRRRGDAVVGASLRDQNHKQTVKML 260

RESULT 13
US-09-077-674-8
; Sequence 8, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-8

Query Match 31.3%; Score 639.5; DB 4; Length 361;
Best Local Similarity 47.0%; Pred. No. 2e-45;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

QY 19 WPAIPPCD---ERRCSPPGALVPVTAVALCLFVGVSGNVVTVMLIGRYDRMTTNL 75
DB 16 WDASPGNDSLGDELQLFPAPLAGVATATCVAFVVGVIAGNLLTMLVVSFRFRLTTNL 75
QY 76 YLGSMVSDLLILGLPDLVRLWRSRPMVPGPLCLSLYVGGCTYATLLHMTALSVE 135
DB 76 YLSMAFSDLLIFLCMPDLVRLWQYRPWNGDILLCKLFQVSESCYATVLTITALSVE 135
QY 136 RYLAICRPLARLVTRRRVRLALIAVLWAVALLSAGPFLVGVQDPGIVVPGVGLNGTA 195
DB 136 RYFAICFPLRAKVTVTKGRVKLVIFVIAVAFCSAGPIFVLVGEHE-----NGT- 185
QY 196 RIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRCRPS--PAQLGALRVMLVTT 253
DB 186 -----DP--W-----DTNECRPTEFAVRSGLLTVMVWVSS 213
QY 254 AYFPLPCLSLYGLIGRLWSSRRLRGLRPAASGRGRHQTQTVRL 300
DB 214 IFPLPVLTVLSLIGRLWRRRRRGDAVVGASLRDQNHQTVKML 260

RESULT 14
US-09-170-496D-210
; Sequence 210, Application US/09170496D
; Patent No. 6553339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6553339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-170-496D-210

Query Match 31.2%; Score 637; DB 4; Length 366;
Best Local Similarity 44.8%; Pred. No. 3.3e-45;
Matches 137; Conservative 43; Mismatches 74; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWPAIPPCD---ERRCSPPGALVPVTAVALCLFVGVSGN 56
DB 2 WNAIPSEEPGNLTLDLADWDAGFNGDNLGDELQLFPAPLAGVATATCVAFVVGVIAGN 61

```

```

QY 57 VVTMLIGRYDRMTTNLYLGSMAVSDLLILGLPDLVRLWRSRPMVPGPLCLSLY 116
DB 62 LLTMLVVSFRFRLTTTNLYLSSMAFSDLLIFLCMPDLVRLWQYRPWNGDILLCKLFQF 121
QY 117 VGGCTYATLLHMTALSVERYLAI CRPLARLVTRRRVRLALIAVLWAVALLSAGPFL 176
DB 122 VSSCTYATVLTITALSVERYFAICFPLRAKVTVTKGRVKLVIFVIAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVVPGVGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECR 236
DB 182 VGVGEH-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFPLCLSLYGLIGRLWSSRRLRGLRPAASGRGRH 294
DB 200 PTEFAVRSGLLTVMVWVSSIFPLPVLVCLTVLSLIGRLWRRRRRGDAVVGASLRDQNHK 259
QY 295 QTVRVL 300
DB 260 QTKKML 265

RESULT 15
US-09-077-675A-7
; Sequence 7, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-7

Query Match 27.1%; Score 554; DB 3; Length 302;
Best Local Similarity 47.7%; Pred. No. 2e-38;
Matches 116; Conservative 35; Mismatches 48; Indels 44; Gaps 5;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:23:05 ; Search time 141.727 Seconds
(without alignments)
546.456 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDGPEAREPPMP.....WQNHLKHGRFADVDLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues
Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1581	77.4	412	12	US-10-318-661-28
2	1581	77.4	412	12	US-10-206-677-2
3	1581	77.4	412	12	US-10-225-567A-473
4	1581	77.4	412	12	US-10-290-078-15
5	1575	77.1	412	12	US-10-417-820A-130
6	1575	77.1	412	12	US-10-417-820A-150
7	649	31.8	289	12	US-10-303-204A-10
8	649	31.8	289	12	US-10-225-567A-140
9	643	31.5	366	12	US-10-303-204A-13
10	643	31.5	366	12	US-10-251-385-88
11	641	31.4	289	12	US-10-303-204A-5
12	641	31.4	364	12	US-10-303-204A-16
13	639.5	31.3	353	12	US-10-303-204A-3
14	639.5	31.3	361	12	US-10-303-204A-8
15	637	31.2	366	15	US-10-251-385-210

16	554	27.1	302	12	US-10-303-204A-7	Sequence 7, Appli
17	552	27.0	302	12	US-10-303-204A-2	Sequence 2, Appli
18	454	22.2	271	12	US-10-303-204A-12	Sequence 12, Appl
19	355.5	17.4	445	12	US-10-240-145-53	Sequence 53, Appl
20	355.5	17.4	445	12	US-10-240-145-139	Sequence 139, Appl
21	354.5	17.4	403	12	US-10-353-690-10	Sequence 10, Appl
22	354.5	17.4	403	15	US-10-251-385-114	Sequence 114, Appl
23	354.5	17.4	403	15	US-10-225-567A-540	Sequence 540, Appl
24	354.5	17.4	403	15	US-10-290-078-18	Sequence 18, Appl
25	352.5	17.3	403	15	US-10-251-385-224	Sequence 224, Appl
26	335	16.4	412	15	US-10-225-567A-557	Sequence 557, Appl
27	335	16.4	415	12	US-10-272-983-12	Sequence 12, Appl
28	335	16.4	415	12	US-10-393-807-12	Sequence 12, Appl
29	335	16.4	415	12	US-10-417-820A-12	Sequence 12, Appl
30	330	16.2	418	15	US-10-225-567A-207	Sequence 207, Appl
31	326.5	16.0	418	12	US-09-826-509-535	Sequence 535, Appl
32	322	15.8	249	11	US-09-782-974C-18	Sequence 18, Appl
33	309	15.1	419	9	US-09-804-551B-26	Sequence 26, Appl
34	309	15.1	428	15	US-10-270-333-114	Sequence 114, Appl
35	285.5	14.0	418	12	US-10-369-493-5319	Sequence 5319, Ap
36	274.5	13.4	410	12	US-09-826-509-537	Sequence 537, Appl
37	274.5	13.4	410	15	US-10-225-567A-432	Sequence 432, Appl
38	268.5	13.1	416	12	US-10-205-219-21	Sequence 21, Appl
39	262	12.8	362	11	US-09-992-331-13	Sequence 13, Appl
40	262	12.8	362	11	US-09-964-923A-24	Sequence 24, Appl
41	261	12.8	595	12	US-10-314-076-17	Sequence 17, Appl
42	261	12.8	595	15	US-10-270-333-195	Sequence 195, Appl
43	259.5	12.7	367	9	US-09-823-114-23	Sequence 23, Appl
44	259.5	12.7	367	15	US-10-290-748-23	Sequence 23, Appl
45	255.5	12.5	378	12	US-10-369-493-6848	Sequence 6848, Ap

ALIGNMENTS

RESULT 1

US-10-318-661-28
; Sequence 28, Application US/10318661
; Publication No. US20030167476A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Bruce R.
; TITLE OF INVENTION: Selective Target Cell Activation By
; Expression of A G Protein-Coupled Receptor Activated
; TITLE OF INVENTION: Superiorly By Synthetic Ligand
; FILE REFERENCE: UCAL-049CIP2
; CURRENT APPLICATION NUMBER: US/10/318,661
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 09/341,446
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US97/05334
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: US 08/622,348
; PRIOR FILING DATE: 1996-03-26
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-318-661-28

Query Match	77.4%	Score 1581;	DB 12;	Length 412;
Best Local Similarity	100.0%	Pred. No. 2,4e-118;		
Matches 300;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MGSPWNGSDGPEAREPPMPALPPCDERCSPPFLGALVPVTA	CLCLFVWGS	GNVTV 60
Db	1	MGSPWNGSDGPEAREPPMPALPPCDERCSPPFLGALVPVTA	CLCLFVWGS	GNVTV 60
Qy	61	MLIGRYDRMRTTNLYLGSMAVSDLLILGLPDLRLRSR	WVFGPILCR	LSLVYVGE 120
Db	61	MLIGRYDRMRTTNLYLGSMAVSDLLILGLPDLRLRSR	WVFGPILCR	LSLVYVGE 120

Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
QY 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
Db 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
QY 241 QLGALRVMLVWTTTAYFPFLPCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTQVRVL 300
Db 241 QLGALRVMLVWTTTAYFPFLPCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTQVRVL 300

RESULT 5

US-10-417-820A-130
; Sequence 130, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; PRIOR FILING DATE: 1999-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-130

Query Match 77.1%; Score 1575; DB 12; Length 412;
Best Local Similarity 99.7%; Pred. No. 7.4e-118;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGVSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGVSGNVTV 60
QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120
Db 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
QY 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
Db 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
QY 241 QLGALRVMLVWTTTAYFPFLPCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTQVRVL 300

Db 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
QY 241 QLGALRVMLVWTTTAYFPFLPCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTQVRVL 300
Db 241 QLGALRVMLVWTTTAYFPFLPCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTQVRVL 300

RESULT 6

US-10-417-820A-150
; Sequence 150, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; PRIOR FILING DATE: 1999-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 150
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-150

Query Match 77.1%; Score 1575; DB 12; Length 412;
Best Local Similarity 99.7%; Pred. No. 7.4e-118;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGVSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCLCLFVVGVSGNVTV 60
QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120
Db 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
QY 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
Db 181 QDGGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
QY 241 QLGALRVMLVWTTTAYFPFLPCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTQVRVL 300

Db 241 QLGALRVMLVTTAYFFLPCLSLYGLIGRELWSSRRPLRGPAASGRGRHGRQTKRVL 300
|||||

RESULT 7

US-10-303-204A-10
; Sequence 10, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR FILING DATE: 1996-12-10
; PRIOR FILING DATE: 1996-06-06
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 289
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-10

Query Match 31.8%; Score 649; DB 12; Length 289;
Best Local Similarity 43.2%; Pred. No. 5.2e-44;
Matches 145; Conservative 46; Mismatches 87; Indels 58; Gaps 8;
QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVAVCLCLFVVGSGN 56
Db 2 WNAATPSEEPGNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALFVVGIA 61
QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLY 116
Db 62 LLTLMVVSREPRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPWVFGDLCCKLQF 121
QY 117 VGEQCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLAIVLWAVALLSAGPFLFL 176
Db 122 VSECTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVFVWVAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVSVPLNGTARIASSPPLWLSRAPPPSPGPETAFAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFFLPCLSLYGLIGRELWSSRRPLRGPAASGRGRHGR 294
Db 200 PTEFAVRSGLLTVMVWSSIFFLPVFCITVLYSLIGRKLWRRRGDAVVGASLRDQNHK 259
QY 295 QTVRLKWSRGSKDACIQSAPGTAQTGLPPL 330
Db 260 QTVKML-----GGQRAURLSLAGPILSLCLPLSL 289

RESULT 8

US-10-225-567A-140
; Sequence 140, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.

; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 2000-12-19
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-140

Query Match 31.8%; Score 649; DB 15; Length 289;
Best Local Similarity 43.2%; Pred. No. 5.2e-44;
Matches 145; Conservative 46; Mismatches 87; Indels 58; Gaps 8;
QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVAVCLCLFVVGSGN 56
Db 2 WNAATPSEEPGNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALFVVGIA 61
QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLY 116
Db 62 LLTLMVVSREPRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPWVFGDLCCKLQF 121
QY 117 VGEQCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLAIVLWAVALLSAGPFLFL 176
Db 122 VSECTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVFVWVAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVSVPLNGTARIASSPPLWLSRAPPPSPGPETAFAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFFLPCLSLYGLIGRELWSSRRPLRGPAASGRGRHGR 294
Db 200 PTEFAVRSGLLTVMVWSSIFFLPVFCITVLYSLIGRKLWRRRGDAVVGASLRDQNHK 259
QY 295 QTVRLKWSRGSKDACIQSAPGTAQTGLPPL 330
Db 260 QTVKML-----GGQRAURLSLAGPILSLCLPLSL 289

RESULT 9

US-10-303-204A-13
; Sequence 13, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR FILING DATE: 1996-12-10
; PRIOR FILING DATE: 1996-06-06
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13

```

; LENGTH: 366
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-13

Query Match      31.5%; Score 643; DB 12; Length 366;
Best Local Similarity 45.1%; Pred. No. 2.1e-43;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

Qy 5 WNGSDGPEGA-----REPPMPALPPCD---ERRCSPPLGALVPVTAACLCFVVGSGN 56
Db 2 WNA TPSEEPGNLTLDLDWDASGNSLDELQLLPAPLAGVTATCTVALFVVGAGN 61

Qy 57 VVTVMILIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLMRSRPWFGLLCRLSLY 116
Db 62 LLTMLVSVSRPRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLMQRYPNGLLCKLPQF 121

Qy 117 VGECTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLIALVLAVALLSAGPFL 176
Db 122 VSECTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181

Qy 177 VGVQDPGIVSVPLGNTARIASSPLASPPMLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DTNECR 199

Qy 237 PS--PAQGLARVLMVWTTAYFPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294
Db 200 PTEFAVRSGLLTVMVVSSIFFFPVFCVLTVLYSLIGRKLWRRRRGDVAVVGASLRDQNHK 259

Qy 295 QTVRVL 300
Db 260 QTVKML 265

RESULT 10
US-10-303-204A-5
; Sequence 5, Application US/10303204A
; Publication No. US200301066144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 289
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-303-204A-5

Query Match      31.4%; Score 641; DB 12; Length 289;
Best Local Similarity 43.3%; Pred. No. 2.3e-43;
Matches 142; Conservative 45; Mismatches 77; Indels 64; Gaps 9;

Qy 5 WNGSDGPEGA-----REPPMPALPPCD---ERRCSPPLGALVPVTAACLCFVVGSGN 56
Db 2 WNA TPSEEPGNLTLDLDWDAPPENDSLVEELLPLPTPLLAGVTATCTVALFVVGAGN 61

Qy 57 VVTVMILIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLMRSRPWFGLLCRLSLY 116
Db 62 LLTMLVSVSRPRELRTTNNLYLSSMAFSELLIFLCMPLELFLMQRYPNGLLCKLPQF 121

Qy 117 VGECTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLIALVLAVALLSAGPFL 176
Db 122 VSECTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181

Qy 177 VGVQDPGIVSVPLGNTARIASSPLASPPMLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DTNECR 199

Qy 237 PS--PAQGLARVLMVWTTAYFPFLCLSLYGLIGRELWSSRRPLRGPAASG---RER 291
Db 200 ATFAVRSGLLTVMVVSSVFFPFLVFCVLTVLYSLIGRKLW---RRKRGDAVVGSSLRDQ 256

Qy 292 GHRQTVRVLKWRSGSKDACLQAPPG 319
; : : : :

```


100 PRIOR APPLICATION NUMBER: 60/008,582
101 PRIOR FILING DATE: 1995-12-13
102 NUMBER OF SEQ ID NOS: 16
103 SOFTWARE: FastSeq for Windows Version 4.0
104 SEQ ID NO 8
105 LENGTH: 361
106 TYPE: PRT
107 ORGANISM: homo sapiens
108 US-10-303-204A-8

Query Match 31.2%; Score 639.5; DB 12; Length 361;
Best Local Similarity 47.0%; Pred. No. 3.9e-43;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

QY	DB	122	VSESCYATVLTITALTALSVRYFAICPPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL	181
QY	QY	177	VGVEQDPGIVVPGNLGTARIASSPLASPPPLWLSRAPPPSPGPETAATAAALFSGRECR	236
DB	DB	182	VGVEHE-----NGT-----DP--W-----DTNECR	199
QY	QY	237	PS--PAQLGALRVMLVWTTAYFPPLPCLLSILYGLIGRELWSSRRRLRGPAAASGRGRGHR	294
DB	DB	200	PTEFAVRSGLLTVWVWSSIFPFPLPVFCLTVLYSLIGRKLWRRRRRGDAVVGASLRDQNHK	259
QY	QY	295	QTVRVL 300	
DB	DB	260	QTKKML 265	

Search completed: January 1, 2004, 06:53:29
Job time : 142.727 sec

100 PRIOR APPLICATION NUMBER: 60/008,582
101 PRIOR FILING DATE: 1995-12-13
102 NUMBER OF SEQ ID NOS: 16
103 SOFTWARE: FastSeq for Windows Version 4.0
104 SEQ ID NO 8
105 LENGTH: 361
106 TYPE: PRT
107 ORGANISM: homo sapiens
108 US-10-303-204A-8

Query Match 31.2%; Score 639.5; DB 12; Length 361;
Best Local Similarity 47.0%; Pred. No. 3.9e-43;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

QY	DB	19	WPALPPCD---ERRCSPPFGALVPTAVCLCLFVVGVSGNVVTVMLIGRYRDMRTTNNL	75
DB	DB	16	WDASPGNDSLGDELLQLFPAPLAGVTATCVAFVVGIAGNLLTMLVVSFRRLRTTNNL <td>75</td>	75
QY	QY	76	YLGSMVSDLLILGLPFDLYRLWRSRPWVPGPLLCRLSLYVGEGCTYATLLHMTALSVE <td>135</td>	135
DB	DB	76	YLGSMVSDLLILFLCMLDLVRLWQYRPMNFGDILCKLFQVSESCYATVLTITALSVE <td>135</td>	135
QY	QY	136	RYLAICRPLRVRVRRVRLALVAVALLSAGPFLVGVQDPGIVVPGNLGTA <td>195</td>	195
DB	DB	136	RYFAICPPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVLVGVGHE-----NGT- <td>185</td>	185
QY	QY	196	RIASSPLASSPPLWLSRAPPPSPGPETAATAAALFSRECRPS--PAQLGALRVMLWVTT <td>253</td>	253
DB	DB	186	-----DP--W-----DTNECRPTEFAVRSGLLTVWVWSS <td>213</td>	213
QY	QY	254	AYFPLPCLLSILYGLIGRELWSSRRRLRGPAAASGRGRGHRQTVRVL <td>300</td>	300
DB	DB	214	IFFFLPVFCLTVLYSLIGRKLWRRRRRGDAVVGASLRDQNHKQTVKML <td>260</td>	260

Search completed: January 1, 2004, 06:53:29
Job time : 142.727 sec

RESULT 15
US-10-251-385-210
; Sequence 210, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-210

Query Match 31.2%; Score 637; DB 15; Length 366;
Best Local Similarity 44.8%; Pred. No. 6.2e-43;
Matches 137; Conservative 43; Mismatches 74; Indels 52; Gaps 7;

QY	DB	5	WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFGALVPTAVCLCLFVVGVSGN	56
DB	DB	2	WNATPSEPGFNTLADLDWDASPGNDSLGDELLQLFPAPLAGVTATCVAFVVGIAGN	61
QY	QY	57	VTVMLIGRYRDMRTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWVPGPLLCRLSLY	116
DB	DB	62	LLTMLVVSFRRLRTTNNLYLSSNAFSDLLIFLCMLDLVRLWQYRPMNFGDILCKLFQF	121
QY	QY	117	VGEGCTYATLLHMTALSVEYLAICRPLRVRVRRVRLALVAVALLSAGPFLFL	176

Query Match 31.2%; Score 637; DB 15; Length 366;
Best Local Similarity 44.8%; Pred. No. 6.2e-43;
Matches 137; Conservative 43; Mismatches 74; Indels 52; Gaps 7;

QY	DB	5	WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFGALVPTAVCLCLFVVGVSGN	56
DB	DB	2	WNATPSEPGFNTLADLDWDASPGNDSLGDELLQLFPAPLAGVTATCVAFVVGIAGN	61
QY	QY	57	VTVMLIGRYRDMRTTNNLYLGSMVSDLLILGLPFDLYRLWRSRPWVPGPLLCRLSLY	116
DB	DB	62	LLTMLVVSFRRLRTTNNLYLSSNAFSDLLIFLCMLDLVRLWQYRPMNFGDILCKLFQF	121
QY	QY	117	VGEGCTYATLLHMTALSVEYLAICRPLRVRVRRVRLALVAVALLSAGPFLFL	176

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 24.6692 Seconds
(without alignments)
1504.757 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDGEGAREPPWP.....WQNHHLKHGRFADVLLSVL 386
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	343	16.8	424	2 JH0164	neurotensin recept
2	330	16.2	418	2 S29506	neurotensin recept
3	316	15.5	477	2 JC7913	capa receptor (CGI
4	295.5	14.0	418	2 A88013	protein K10B4.4 [I
5	268.5	13.1	416	2 S68822	neurotensin recept
6	255.5	12.5	378	2 T15816	hypothetical prote
7	253.5	12.4	367	2 I49022	kappa opioid recep
8	253.5	12.4	367	2 JC2421	opioid receptor ho
9	253.5	12.4	367	2 I56520	G protein-coupled
10	253.5	12.4	370	2 S43087	orphan opioid rece
11	250.5	12.3	363	2 I57940	somatostatin recep
12	249.5	12.2	452	2 A46195	cholecystokinin B
13	249.5	12.2	452	2 JC2459	gastrin/cholecysto
14	248.5	12.2	418	2 A46226	somatostatin recep
15	247	12.1	453	2 S32817	gastrin receptor -
16	245.5	12.0	363	2 I57955	somatostatin recep
17	245.5	12.0	364	2 JN0763	somatostatin recep
18	243	11.9	359	2 JC5277	G protein-coupled
19	241.5	11.8	477	1 ORHUB1	beta-1-adrenergic
20	240	11.7	450	2 JQ1614	gastrin receptor -
21	239	11.7	352	2 JE0296	thyrotropin releas
22	237.5	11.6	447	2 A47430	gastrin/cholecysto
23	237.5	11.6	519	2 S17783	tachykinin recepto
24	234.5	11.5	514	2 D56849	dopamine receptor-
25	233.5	11.4	480	2 I53053	beta 1 adrenergic
26	232.5	11.4	379	2 JC6178	serotonin receptor
27	232.5	11.4	450	2 I49481	alpha-2 adrenergic
28	232	11.4	380	2 A55259	kappa opioid recep
29	230	11.3	477	2 T25846	hypothetical prote

30	229.5	11.2	466	2 S36794	beta-1-adrenergic
31	229	11.2	427	2 S50150	gastric CCK-A rece
32	228.5	11.2	387	2 JC5949	galanin receptor 2
33	228	11.2	437	2 I57942	5-hydroxytryptamin
34	228	11.2	450	2 B40392	alpha-2-adrenergic
35	227.5	11.1	428	2 S30508	probable G protein
36	227	11.1	380	2 JC2338	kappa opioid recep
37	227	11.1	418	2 G02953	beta-3-adrenergic
38	226	11.1	398	2 JN0708	thyrotropin-releas
39	225	11.0	444	2 A42685	cholecystokinin re
40	224.5	11.0	411	2 I56444	thyrotrophin-relea
41	224.5	11.0	412	2 S23436	thyrotroberin recep
42	224.5	11.0	428	2 A44021	somatostatin recep
43	224	11.0	483	2 A25896	beta-adrenergic re
44	223.5	10.9	436	2 JC5599	cholecystokinin-A
45	223.5	10.9	450	2 A38316	alpha-2-adrenergic

ALIGNMENTS

RESULT 1

JH0164
neurotensin receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000
C;Accession: JH0164
R;Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A;Title: Structure and functional expression of the cloned rat neurotensin receptor.
A;Reference number: JH0164; MUID:90297956; PMID:1694443
A;Accession: JH0164
A;Molecule type: mRNA
A;Residues: 1-424 <TAN>
C;Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor. This
ter (neuromodulator in the brain and as a hormone) cellular mediator in peripheral tiss
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;65-87/Domain: transmembrane #status predicted <TM1>
F;97-121/Domain: transmembrane #status predicted <TM2>
F;144-165/Domain: transmembrane #status predicted <TM3>
F;189-210/Domain: transmembrane #status predicted <TM4>
F;236-260/Domain: transmembrane #status predicted <TM5>
F;309-330/Domain: transmembrane #status predicted <TM6>
F;348-372/Domain: transmembrane #status predicted <TM7>
F;4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.8%; Score 343; DB 2; Length 424;
Best Local Similarity 33.9%; Pred. No. 4.7e-19;
Matches 82; Conservative 41; Mismatches 71; Indels 48; Gaps 6;

QY	39	VPTAVCLCLFVVGSGNVVTVMLIGR---YRDMRTTNLYLGSMVSDLLI-LLGLPPD	94
DB	65	VLVTALYALFVVGTVGNSVTFTLARKKSLSQSTVHYHLGSLSDLSLLILLAMPVE	124
QY	95	LYR-LWRSPWVPGPLLCRLSLVYVGSGCTVATLHMTALSVERYLAICRPLARVLVTRR	153
DB	125	LYNFIWHPWAGDAGCGGYFLRDACYATATLVNLSVERYLAICHPKATLMRSR	184
QY	154	RVRALIVLWAVALLSAGPFLFLVVGVEQDPGIVSVFGLNGTARIASSPLASSPPLMSRA	213
DB	185	RTKKFISAIWLASALLAIPMLFTMLQNRSGDGTHPG-----GLVCTPIVDT-----	231
QY	214	PPSPSGPETAATAAALFSECRPSPAQLGALRVMLVWVTAYFFL-PPFLCSILYGLIGR	272
DB	232	-----ATVKVVIQVNTFMGFLPPLMVLISILNTVIAN	262
QY	273	EL 274	
DB	263	KL 264	

RESULT 2

```

Db      146 AFISEACTYVSVFTIVAFSMERFVAICHPLHLVAMVGPKRAIRITALTALWVFSISAIPF- 204
      : : ||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy      175 FLVGVQDPGIVSVVPGVGLNGTARIASSPIASSPPLWLSRAPPPSPGPETAFAAALFSRE 234
      ||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      205 -----GLSDIQLYNPLDHS-----RIEESAF----- 227
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy      235 CRSPSAQLGALRVLMTWTAYFFLPFLCLSTLYGLIGRELWSSRRRLRG--PAASGRERG 292
      ||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      228 CSMSPKIVNEIPVEVSFCIPFVIFPMILIILLYGRMGAKIRSRTNQKLGVOQGTNNRETR 287
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy      293 HRQ-----TVRVL 300
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      288 NSQMRKKTIVRML 300
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 4
A88013
protein K10B4.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88013
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigation
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Proj
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103,
A:Accession: A88013
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <STO>
A:Cross-references: GB:chr II; PIDN:AB71009.1; PID:92429475; GSPDB:GN00020; C
C:Genetics:
A:Gene: K10B4.4
A:Map position: 2

Query Match          14.0%; Score 285.5; DB 2; Length 418;
Best Local Similarity 27.8%; Pred. No. 1.3e-14;
Matches 79; Conservative 60; Mismatches 114; Indels 31; Gaps 9

Qy      29 RCGSPFPGALVPTAVCLCLFVVGSGVNVTVMLIGRYRDMRTTNTLYLGSMAVSDLL-I 87
      ||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      24 RCQ--SAGIVPTVIYGTIFLLGLFGNICTCIVIAANKSMHNPNTNYLFLSLAVSDIIL 81
      ||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy      88 LLGLPFDLYR-LWRSRPVWGFLLCRLSLYVGECTYATLHMTALSVERYLAICRPLRA 146
      ||||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      82 ILGLPMFYQSLDYSYPRFSEGIKARAFIIEFTSYASIMIICCFSPERMLAICHPLRS 141
      ||||| : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy      147 RVLVTRRRRALIAVLAVALLSAGPFLFLVGVQDP--GISVVPGLNGTARIASSPLA 203
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      142 KIFSTLRANVLIITAWTISFVCALPTAFIVQINKPLPEBAKQYQWTKVKSFFVAVGVLN 201
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Qy      204 SSPPLWLSRAPPPSPGPETAFAAALFSREC-----RPSPAQLGALRVLMTWTYTFPL 258
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      202 -----NRIFPV-----TDGIFVLHTEFCAMNQSRPDQKM-----LIIFAFTVFFVI 244
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

Qy      259 PFLCLSTLYGLIGRELWSSRRRLRG--AASGRERGHQTVRVL 300
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      245 PAIAIVMYAHIAVQLESSEIDLKGDKMVKKRRKNSNRTVLKML 288
      : : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 5
S68822
neurotensin receptor 2, levocabastine-sensitive - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
C:Accession: S68822
R:Chalon, P.; Vita, N.; Kaghad, M.; Guillemot, M.; Bonnin, J.; Delpach, B.; le
FEBS Lett. 386, 91-94, 1996
A:Title: Molecular cloning of a levocabastine-sensitive neurotensin binding si
A:Reference number: S68822; MUID:96228041; PMID:8647296

```

A;Accession: S69822
A;Molecule type: mRNA
A;Residues: 1-416 <CHA>
A;Cross-references: GB:X97121; NID:g1483579; PIDN:CAA65787.1; PID:g1483580
A;Experimental source: hypothalamus
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein
F;33-59/Domain: transmembrane #status predicted <TM1>
F;70-91/Domain: transmembrane #status predicted <TM2>
F;110-131/Domain: transmembrane #status predicted <TM3>
F;155-175/Domain: transmembrane #status predicted <TM4>
F;204-230/Domain: transmembrane #status predicted <TM5>
F;296-315/Domain: transmembrane #status predicted <TM6>
F;335-361/Domain: transmembrane #status predicted <TM7>
Query Match 13.1%; Score 268.5; DB 2; Length 416;
Best Local Similarity 37.9%; Pred. No. 2.6e-13;
Matches 66; Conservative 33; Mismatches 66; Indels 9; Gaps 4;
QY 18 PWPALP-----PCDERCCSPFPLGALVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMRT 71
DB 6 PWPSPSPSAGLSLEARLGVDTLWAKVLTALYSYLIFAFGTAGNALSVMHVKARAGRP 65
QY 72 -TTNLYLGSMAVSDLLILL-GLPFDLYR-IWRSRPWVFGPLCLSLYVGEGETYATLLH 128
DB 66 GLRUYHVLVSALLSALLLLVSMPEMLYNFVMSHPWVFGDLGCRGYFVRELCAATVLS 125
QY 129 MTSLVRYLAICRPLRARVLTTRRRVRLIALIWAVALLSAGPFLPLVGVQED 182
DB 126 VASLSAERCLAVCQPLRRRLRLPRTRRLSLVVASLGLALPMAVIMQKH 179
RESULT 6
T15816
hypothetical protein C48C5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C;Accession: T15816
R;Favella, A.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C48C5.
A;Reference number: Z18410
A;Accession: T15816
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-378 <FAV>
A;Cross-references: EMBL:U39994; NID:g1055102; PID:g1055105; PIDN:AAB37017.1; GSPDB:GN00
A;Experimental source: strain Bristol N2; clone C48C5
C;Genetics:
A;Gene: CESP:C48C5.1
A;Map position: X
A;Introns: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1
C;Superfamily: adenosine receptor A1
Query Match 12.5%; Score 255.5; DB 2; Length 378;
Best Local Similarity 27.0%; Pred. No. 2.4e-12;
Matches 66; Conservative 49; Mismatches 96; Indels 33; Gaps 6;
QY 38 LVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMRTTNNLYLGSMAVSDLLIL-LGLPDL 96
DB 51 LYKVYALYIFLFGVIGNTTCLVMKKHMKTHASMLNLAVALSDVLTLCVGLPPEVM 110
QY 97 RLWRSRPWVFGPLCLSLYVGEGETYATLLHMTALSVRYLAICRPLRARVLTTRRV 155
DB 111 MNWQYWPFPDYICNLKALIAETSSVSLITLIFAIERYVAVCHFLMKVQPPKRN 170
QY 156 RLIALIWAVALLSAGPFLPLVGVQEDPGISVVPGLNGTARIASSPPLWLSRAP 215
DB 171 GTIIGTWISILCAMP--FAIHRADYIMKSWGTNDNRIPVKSCKWC----- 217
QY 216 PSPSPGPETAALFSPRCRPSPAQLGALRVML-WYTTAYFFLPFLCLSLYGLIGREL 274
DB 218 -----IAVMF-----EPKLASTKILFHFSAIAFFALPLFTIVILYARIACK 260

QY 275 WSSR 278
DB 261 SSNR 264
RESULT 7
I49022
kappa opioid receptor 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C;Accession: I49022
R;Pan, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brooke, A.I.; De
Mol. Pharmacol. 47, 1180-1188, 1995
A;Title: Cloning and functional characterization through antisense mapping of a kappa 3-
A;Reference number: I49022; MUID:95327076; PMID:7603458
A;Accession: I49022
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-367 <RES>
A;Cross-references: EMBL:U09421; NID:g551484; PIDN:AAA81333.1; PID:g551485
C;Superfamily: vertebrate rhodopsin
Query Match 12.4%; Score 253.5; DB 2; Length 367;
Best Local Similarity 26.8%; Pred. No. 3.3e-12;
Matches 72; Conservative 47; Mismatches 95; Indels 57; Gaps 6;
QY 34 PLGALVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMRTTNNLYLGSMAVSDLLILGLPF 93
DB 44 PLGLKVTIVGLYLAVCIGGLGCLVMYILRHTKMTATNIYIFNLALADTLVLLTLP 103
QY 94 ---DLYRLWRSRPWVFGPLCLSLYVGEGETYATLLHMTALSVRYLAICRPLRARV 150
DB 104 QGTDILLGP-----WPFGNALCKTVIAIDVYVNMFTSTFTLTAMSDRVYVAICHDRLDVR 159
QY 151 TRRVVATLAVLWAVALLSAGPFLPLVGVQEDPGISVVPGLNGTARIASSPPLW 210
DB 160 TSSKAQAVNAIWAALASV-----VGPVVAIMGSAQVDEBEIEC-----L 198
QY 211 SRAPPPSPSPGPETAALFSPRCRPSPAQLGALRVMLWYTTAYFFLPFLCLSLYGLI 270
DB 199 VEIPAPQDYWGVPVFAICIFLFS-----FIIPVLIISVCYSLM 235
QY 271 GRELWSSRRPLRGPAASGRGRGHRTQVRVLR 301
DB 236 IRLRGVR-----LLSGSREKDRNLRRITR 260
RESULT 8
JC2421
opioid receptor homolog, MOR-C - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 24-Nov-1999
C;Accession: JC2421; I49122
R;Nishi, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
A;Title: Structure and chromosomal mapping of genes for the mouse kappa-opioid receptor
A;Reference number: JC2421; MUID:95100967; PMID:7802669
A;Accession: JC2421
A;Molecule type: mRNA
A;Residues: 1-367 <NIS>
A;Cross-references: DDBJ:D31663
R;Halford, W.P.; Gebhardt, B.M.; Carr, D.J.
J. Neuroimmunol. 59, 91-101, 1995
A;Title: Functional role and sequence analysis of a lymphocyte orphan opioid receptor.
A;Reference number: I49122; MUID:95318231; PMID:7797625
A;Accession: I49122
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-357 <RES>
A;Cross-references: EMBL:U14165; NID:g540092; PIDN:AAA87899.1; PID:g540093
C;Genetics:
A;Gene: MOR-c; OOR


```
Query Match      12.3%   Score 250.5; DB 2; Length 363;
Best Local Similarity 26.7%; Pred. No. 5.5e-12;
Matches 76; Conservative 43; Mismatches 99; Indels 67; Gaps 8;

Qy 5 WNGSDGPEGAREPPWALPCCDERRCFFPLGA---LVPVTAVCLCLFVVGVSNGVVTVM 61
12 WNASAASSGNHN--WSLVG-----SASPMPGARVLVPLYLLVC--TVGLSGNTIAIV 60
Db
```

Qy	48	LFVVGVSGNVVTVMLIGRYDRMRTTNNLYLGSMAVSDLLILG-LPFDDYLWRSRP----	103
Db	64	IFLMSGVGNVLIIVVLGLSRRLRTVNAFLSLAVSDLLLAACMPFTLL-----PNLM	117
Qy	104	--WVFGPLLCL-LSLYVCEGCTYATLLHWTALSVERYLAICRPLBARVLVTRRRVRLIA	160
Db	118	GTFIFGTWICKAISYLMGVSVSVST-LNLVAIALERYGAICRPLQARVQTRSHAARVTL	176
Qy	161	VLMAVALLSAGPFLFLVGVGEODPGISVVPGLNGTARIASSPLASSPPLWLRAPPPSPS	220
Db	177	ATWLUSGLLMVPP-----PVVTWV-----QPV	198
Qy	221	GPETAEEAALFSRECRPSPAQALGARVMLWTVTAYFFLPFLCLSLYGLIGRELW----	275
Db	199	GPRVLQCHRW-----PSARVQQTWSVLLLL--LFFPIGVVIAYAGLISRELYLGLHF	251
Qy	276	-----SRRRPLRGPAAG-----REGRHQTIVRLKWRSGSKDACLOSPPGT	320
Db	252	DGENDSETQSRARNQGGPLPGCAAGPVPVHONGCRPVTSVAGE-----DSDGCCVQLPRSR	306
Qy	321	AQTGLPPLLAQLWAPLPAPFP	342
Db	307	LE-----MTTLTTPTEGPVP	321

```

RESULT 13
JC2459
gasrin/cholecystokinin B receptor - rabbit
C:Species: Cryptolagus cuniculus (domestic rabbit)
C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 20-Apr-2000
C:Accession: JC2459
R:Blandizzi, C.; Song, I.; Yamada, T.

```

Biochem. Biophys. Res. Commun. 202, 947-953, 1994
 A;Title: Molecular cloning and structural analysis of the rabbit gastrin/CKKB receptor cDNA
 A;Reference number: JC2459; MUID:94324990; PMID:8048969
 A;Accession: JC2459
 A;Molecule type: mRNA
 A;Residues: 1-452 <BLA>
 A;Cross-references: GB:L31548; NID:G495663; PIDN:AAA31194.1; PID:G495665
 C;Genetics:
 A;Introns: 49/1; 133/1; 216/2; 273/1
 C;Superfamily: neurokinin 1 receptor
 C;Keywords: receptor; transmembrane protein
 F;56-79/Domain: transmembrane #status predicted <TM1>
 F;85-104/Domain: transmembrane #status predicted <TM2>
 F;130-149/Domain: transmembrane #status predicted <TM3>
 F;169-187/Domain: transmembrane #status predicted <TM4>
 F;217-237/Domain: transmembrane #status predicted <TM5>
 F;339-359/Domain: transmembrane #status predicted <TM6>
 F;381-400/Domain: transmembrane #status predicted <TM7>

Query Match 12.2%; Score 249.5; DB 2; Length 452;
 Best Local Similarity 25.2%; Pred. No. 8.3e-12;
 Matches 86; Conservative 50; Mismatches 122; Indels 83; Gaps 10;

QY 48 LFVVGSGNVVTVMLIGRYEDMTTNNLYLGSMVAVSDLLILGL-LPFDLYRLWRSRP--- 103
 Db 62 IFLMSVGGNLIIVLVGLSRRLRTVTNAPFLSLAVSDLLAVACMPETLL-----PNLM 115

QY 104 --WVFGPCLRLSLYVGEGCTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAV 161
 Db 116 GTFIFGTICKAVSLGMSVSVSTLSVAIALERYSAICRPLQARVQWTRSHAARVILA 175

QY 162 LMAVALLSAGPFLVGVQEDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPSG 221
 Db 176 TWLSGLLMVPIPVYTAQ-----PVG 197

QY 222 PETAEAAALFSRECRPSAQLGALRYMLWVTTAYFFLPFLCLSLYGLIGRELWSSRR-- 279
 Db 198 PRVLQCVHRM-----FSARVQRTWSVLLLL--LLFFVGVVMAVAYGLISRELYLGLRFD 250

QY 280 -----PLRGPAASGRERGHQRTVRLKRSRSG--DACLOSAPPTAQ 323
 Db 251 SDSDESQSRVRQGGPLGGAAPGPV---HONGRCRPEAGLEDGCGYVQLPRSR--- 304

QY 324 LGPLPALLAOLWAPLPAPFISIPASTRRGGSGIYNLLVAL 364
 Db 305 --PALELSALTAPISGPGPRPAQAKLLAKRVRLVI 343

RESULT 14
 A46226
 C;Species: Homo sapiens (man)
 C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
 C;Accession: A46226; S32501
 R;Yamada, Y.; Reisine, T.; Law, S.F.; Ihara, Y.; Kubota, A.; Kagimoto, S.; Seino, M.; Seino, M.; Endocrinol. 6, 2136-2142, 1992
 A;Title: Somatostatin receptors, an expanding gene family: cloning and functional characterization
 A;Reference number: A46226; MUID:93149123; PMID:1337145
 A;Accession: A46226
 A;Molecule type: DNA
 A;Residues: 1-418 <YAM>
 A;Cross-references: GB:M96738; NID:G338498; PIDN:AAA60592.1; PID:G338499
 A;Note: sequence extracted from NCBI backbone (NCBI:123685, NCBI:123690)
 R;Corrness, J.D.; Demchynshyn, L.L.; Seeman, P.; van Tol, H.H.M.; Srikant, C.B.; Kent, G.; FEBS Lett. 321, 279-284, 1993
 A;Title: A human somatostatin receptor (SSTR3), located on chromosome 22, displays preference for the [D-Trp] form of somatostatin
 A;Reference number: S32501; MUID:93238970; PMID:8097479
 A;Accession: S32501
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-418 <COR>
 C;Genetics:
 A;Gene: GDB:SSTR3

A;Cross-references: GDB:134187; OMIM:182453
 A;Map position: 22q13.1-22q13.1
 A;Introns: #status absent
 C;Superfamily: vertebrate rhodopsin
 C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
 F;44-70/Domain: transmembrane #status predicted <TM1>
 F;81-106/Domain: transmembrane #status predicted <TM2>
 F;118-139/Domain: transmembrane #status predicted <TM3>
 F;159-181/Domain: transmembrane #status predicted <TM4>
 F;203-233/Domain: transmembrane #status predicted <TM5>
 F;255-282/Domain: transmembrane #status predicted <TM6>
 F;289-316/Domain: transmembrane #status predicted <TM7>
 F;17-30/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;116-191/Disulfide bonds: #status predicted
 F;151-251,317,332/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 F;251/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
 F;256/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predicted
 F;412/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 12.2%; Score 248.5; DB 2; Length 418;
 Best Local Similarity 23.7%; Pred. No. 9.1e-12;
 Matches 88; Conservative 51; Mismatches 133; Indels 99; Gaps 11;

QY 11 PGAREPPPPALPPCDERRCSPFPL-----GALVPVTAVCLCLFVVGVSQNVVTVMLIGR 65
 Db 15 PENA-SSAWPPDPATLGNVSAGSPAGLAVSGVLIPLVAVVC--VVGLLGNLSLVIYVLR 71

QY 66 YDMRTTNNLYLGSMVAVSDLLILGLPDLRLMRSPVWFGPCLRLSLYVGEGCTYAT 125
 Db 72 HTASPSVTNNVITNLALADELFMLGLFF-LAAQNALSYWPFGLMCLVMAVDGINQFTS 130

QY 126 LLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLAVALLSAGPFLFLVGVQEDPGI 185
 Db 131 IFCLTVMSVDRLAVVHPTRSAWRTAPVARTVSAVAVVLPVVFVSGV----- 184

QY 186 SVVPGINGTARIASSPLASSPPLMLSRAPPSPGPETAEEAALFSRRCR---PSPAQL 242
 Db 185 -----PRGMS-----CHMQWPEPAAA 201

QY 243 GALRVMLWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRERGHQRTVRLK 302
 Db 202 WRAGFIITAYALGFGFGLVLCYLLIIVKRSAGRVWAPSCQRRRSRRVTRM-- 259

QY 303 WSRGSKDACLOSAPPTAQTLGPLLAQLWAP-----LPAPFISIPASTRRGG 354
 Db 260 -----VAVVALFVLCWMPFYVLNIVNVVCLP-EEFAF----- 291

QY 355 SGIYNLLVALP 365
 Db 292 FGLYFLVVALP 302

RESULT 15
 S32817
 C;Species: Canis lupus familiaris (dog)
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Apr-2000
 R;Kopin, A.S.; Lee, Y.M.; McBride, E.W.; Miller, L.J.; Lu, M.; Lin, H.Y.; Kolakowski Jr., Proc. Natl. Acad. Sci. U.S.A. 89, 3605-3609, 1992
 A;Title: Expression cloning and characterization of the canine parietal cell gastrin receptor
 A;Reference number: S32817; MUID:92228835; PMID:1373504
 A;Accession: S32817
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-453 <KOP>
 A;Cross-references: EMBL:M87834; NID:G163956; PIDN:AAA30847.1; PID:G163957
 C;Superfamily: neurokinin 1 receptor
 C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 12.1%; Score 247; DB 2; Length 453;
 Best Local Similarity 25.9%; Pred. No. 1.3e-11;
 Matches 97; Conservative 50; Mismatches 110; Indels 118; Gaps 13;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 13.5439 Seconds
(without alignments)
1340.261 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDGPEGAREPPWP.....WQNLHKGHGFADVLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	77.4	412	1	MTLR_HUMAN
2	643	31.5	366	1	GHSR_HUMAN
3	641	31.4	364	1	GHSR_RAT
4	641	31.4	366	1	GHSR_PIG
5	614	30.1	257	1	GHSR_MOUSE
6	343	16.8	424	1	NTR1_RAT
7	340	16.6	424	1	NTR1_MOUSE
8	330	16.2	418	1	NTR1_HUMAN
9	274.5	13.4	410	1	NTR2_HUMAN
10	268.5	13.1	416	1	NTR2_RAT
11	268	13.1	454	1	GASR_BOVIN
12	267.5	13.1	370	1	OPRX_CAVPO
13	266	13.0	417	1	NTR2_MOUSE
14	262	12.8	362	1	SSR5_MOUSE
15	254.5	12.5	370	1	OPRX_PIG
16	253.5	12.4	367	1	OPRX_MOUSE
17	253.5	12.4	367	1	OPRX_RAT
18	253.5	12.4	370	1	OPRX_HUMAN
19	253	12.4	453	1	CCNR_XENLA
20	251.5	12.3	473	1	BIAR_CANFA
21	251	12.3	453	1	GASR_MOUSE
22	250.5	12.3	363	1	SSR5_RAT
23	249.5	12.2	452	1	GASR_RABIT
24	249.5	12.2	452	1	GASR_RAT
25	248.5	12.2	360	1	GP25_HUMAN
26	248.5	12.2	418	1	SSR3_HUMAN
27	247.5	12.1	452	1	A2AA_BOVIN
28	247.5	12.1	453	1	GP39_HUMAN
29	247	12.1	453	1	GASR_CANFA
30	245.5	12.0	364	1	SSR5_HUMAN
31	243	11.9	368	1	GALT_HUMAN
32	241.5	11.8	477	1	BIAR_HUMAN
33	241	11.8	370	1	GALT_RAT

34	240	11.7	450	1	GASR_PRANA
35	239	11.7	372	1	GALS_RAT
36	239	11.7	405	1	B3AR_CANFA
37	237.5	11.6	371	1	GALS_MOUSE
38	237.5	11.6	447	1	GASR_HUMAN
39	237.5	11.6	519	1	TLR2_DROME
40	236	11.6	289	1	SSRL_FUGRU
41	233.5	11.4	480	1	BIAR_MACMU
42	233	11.4	395	1	TRPR_CHICK
43	232.5	11.4	351	1	B3AR_CAVPO
44	232.5	11.4	379	1	GRE2_BALAM
45	232.5	11.4	450	1	A2AA_MOUSE

ALIGNMENTS

RESULT 1	MTLR_HUMAN	STANDARD;	PRT;	412 AA.
AC	O43193;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Motilin receptor (G protein-coupled receptor GPR38)			
GN	GPR38 OR MTLR1 OR MTLR			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]				
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RX	MEDLINE=98110578; PubMed=9441746;			
RA	McKee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,			
RA	Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;			
RT	"Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors."			
RT	Genomics 46:426-434(1997).			
RL				
[2]				
RP	SEQUENCE FROM N.A. (ISOFORMS A AND B).			
RX	MEDLINE=99316084; PubMed=10381885;			
RA	Feighner S.D., Tan C.P., McKee K.K., Palyha O.C., Hreniuk D.L.,			
RA	Pong S.-S., Austin C.P., Figueroa D., MacNeil D., Cascieri M.A.,			
RA	Nargund R., Bakshi R., Abramovitz M., Stocco R., Kargman S.,			
RA	O'Neill G., van Der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,			
RA	Howard A.D.;			
RT	"Receptor for motilin identified in the human gastrointestinal system."			
RT	Science 284:2184-2188(1999).			
RL				
[3]				
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RA	Wall M.;			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
[4]				
RP	FUNCTION.			
RX	MEDLINE=21219832; PubMed=11322507;			
RA	Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,			
RA	Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;			
RT	"Growth hormone secretagogue receptor family members and ligands."			
RL	Endocrine 14:9-14(2001).			
CC	-1- FUNCTION: Receptor for motilin.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=A;			
CC	Isoid=O43193-1; Sequence=Displayed;			
CC	Name=B;			
CC	Isoid=O43193-2; Sequence=VSP_001894;			
CC	-1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE MARROW.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF034632; AAC26081.1; -;
 ENBL; AL137000; CAC19107.1; -;
 Genew; HGNC:4495; GPR38.
 MIM; 602885; -;
 GO; GO:0005887; C:integral to plasma membrane; TAS.
 GO; GO:0007586; P:digestion; TAS.
 GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
 InterPro; IPR000276; GPCR_Rhodopsn.
 Pfam; PF00001; 7tm 1; 1.
 PRINTS; PR00237; GPCR_Rhodopsn.
 PROSITE; PS00237; G PROTEIN RECP FL 1; 1.
 PROSITE; PS0262; G PROTEIN RECP FL 2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein;
 Alternative splicing.
 DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 36 56 1 (POTENTIAL).
 DOMAIN 57 74 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 75 94 2 (POTENTIAL).
 DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 113 134 3 (POTENTIAL).
 DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 158 178 4 (POTENTIAL).
 DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 247 270 5 (POTENTIAL).
 DOMAIN 271 298 6 (POTENTIAL).
 TRANSMEM 299 320 7 (POTENTIAL).
 DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 335 358 7 (POTENTIAL).
 DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 411 235 BY SIMILARITY.
 N-LINKED (GLNAC. . .) (POTENTIAL).
 N-LINKED (GLNAC. . .) (POTENTIAL).
 LVLVLAFLICWLPFHVRIIYNTEDSRMVFQYFNVAL
 QYLVASINPIILNLSIKRYAAAFKLLARKSRGRGHR
 SRTAGEVAGDTGDTVGTETSNVKTWG -> RKWSRRG
 SKDACLQSPGPTAOTLGLPLLAOLWAPLAPFPISIPAS
 TRGGSGGIYNALVALPRWQNLHKKHGFADVVLSVL
 (in isoform B).
 /FTID=VSP_001894.
 SQ SEQUENCE 412 AA; 45344 MW; C13FF6165012DEF3 CRC64;
 Query Match 77.4%; Score 1581; DB 1; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.6e-101; Mismatches 0; Indels 0; Gaps 0;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGEGARPPWPPALPCDERRCPSPFLGALVPTAVCLCLFVVGSGNVTV 60
 DB 1 MGSPWNGSDGEGARPPWPPALPCDERRCPSPFLGALVPTAVCLCLFVVGSGNVTV 60
 QY 61 MLIGRYDRMTTNNLYLGSMVAVSDLLILGLPDLRLMRSRWVPGPLLCRLSLVVGEG 120
 DB 61 MLIGRYDRMTTNNLYLGSMVAVSDLLILGLPDLRLMRSRWVPGPLLCRLSLVVGEG 120
 QY 121 CTYATLLHMTALSVRYLAICRPLRAVLVTRRRVALIAVLWALLSAGPFLVLGVGE 180
 DB 121 CTYATLLHMTALSVRYLAICRPLRAVLVTRRRVALIAVLWALLSAGPFLVLGVGE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPSPGPTAAALFSSRECPSPA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPSPGPTAAALFSSRECPSPA 240
 QY 241 QLGLARVLMVWTTAYFPFLPCLLSILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300
 DB 241 QLGLARVLMVWTTAYFPFLPCLLSILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300

RESULT 2

GHRSR_HUMAN STANDARD; PRT; 366 AA.
 AC Q92847; Q92848; Q96R07;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).
 DE GHRS.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
 RC TISSUE=Pituitary;
 RX MEDLINE=96337998; PubMed=868086;
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P., Liberato P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L., Palyha O.C., Anderson J.J., Parese P.S., Diaz C., Chou M., Liu K.K., McKee K.K., Pong S.-S., Chung L.-Y., Elbrecht A., Dashkevich M., Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G., Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K., Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
 RT "A receptor in pituitary and hypothalamus that functions in growth hormone release.";
 RT Science 273:974-977(1996).
 RL [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
 RX MEDLINE=21255649; PubMed=11356716;
 RA Petersenn S., Rasch A.C., Penschorn M., Beil F.U., Schulte H.M.;
 RT "Genomic structure and transcriptional regulation of the human growth hormone secretagogue receptor.";
 RT Endocrinology 142:2649-2659(2001).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=21219832; PubMed=11322507;
 RA Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
 RT "Growth hormone secretagogue receptor family members and ligands.";
 RL Endocrine 14:9-14(2001).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=20067959; PubMed=10604470;
 RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
 RT "Ghrelin is a growth-hormone-releasing acylated peptide from stomach.";
 RT Nature 402:656-660(1999).
 RL Nature 402:656-660(1999).
 CC -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1A;
 CC IsoId=Q92847-1; Sequence=Displayed;
 CC Name=1B;
 CC IsoId=Q92847-2; Sequence=VSP_001916, VSP_001917;
 CC -!- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).


```
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 364 CYTOPLASMIC (POTENTIAL).
FT DISULFID 115 197 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 40963 MW; DCBF59B061EE9 CRC64;

Query Match 31.4%; Score 641; DB 1; Length 364;
Best Local Similarity 45.8%; Pred. No. 3.6e-37;
Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

Qy 5 WNGSDGPEGAREP-----PWPALPPCD-----ERRCSPFPLGALVPTAVCLCLFVVGSG 55
Db 2 WNAT--PSEPEPNVTLDDWDASGPNDSLPPDELLPLFPAPLLAGVATCATCVLFVVGISG 59
Qy 56 NVVTMLIGRYDRMTTNNLYLSGMVSDLLILGLPDLVRLWRSPPWFGPLCLRLSL 115
Db 60 NLLTMLVVSFRPRELRTTNNLYLSMAFSDLLIFLCMLDLVRLWQYRPNFGDLLCKLFQ 119
Qy 116 YVGGCTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVILVWAVAFCSAGPIFV 179
Qy 176 LVGVEQDPSVWVPLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSREC 235
Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
Qy 236 RPS--PAQLGALRVMLWVTYATFELPCLSLILYGLIGRELWSSRRPLRGPAAAG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVFCLTVLYSLIGRKLW--RR--RGDAVAGSLRD 253
Qy 291 RGHRTQVRVL 300
Db 254 QNHQTVKML 263

RESULT 4
ID_GHSR_PIG STANDARD; PRT; 366 AA.
AC Q95254; Q95255;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHR) (Ghrelin receptor).
DE GHSR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RC STRAIN=Yorkshire; TISSUE=pituitary;
RX MEDLINE=96337998; PubMed=868086;
RA Howard A.D., Feighner S.D., Cully D.F., Hamelin M., Hreniuk D.L., Liu K.K.,
RA Liberator P.A., Rosenblum C.I., Pares P.S., Diaz C., Chou M., Liu K.K.,
RA Palyla O.C., Anderson J., Pares P.S., Chang L.-Y., Elbrecht A., Dashkevich M.,
RA McKee K.K., Pong S.-S., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,
RA Heavens R., Rigby M., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RA "A receptor in pituitary and hypothalamus that functions in growth
RT hormone release."
RT Science 273:974-977(1996).
RL
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1A;
```

Query Match

FT TRANSMEM 144 165 3 (POTENTIAL).
 FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 189 210 4 (POTENTIAL).
 FT DOMAIN 211 235 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 236 260 5 (POTENTIAL).
 FT DOMAIN 261 308 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 309 330 6 (POTENTIAL).
 FT DOMAIN 331 348 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 349 372 7 (POTENTIAL).
 FT DOMAIN 373 424 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 142 225 BY SIMILARITY.
 FT LIPID 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA; 47054 MW; A9C2P7EAF89BCD3 CRC64;

Query Match 16.8%; Score 343; DB 1; Length 424;
 Best Local Similarity 33.9%; Pred. No. 1.1e-16;
 Matches 82; Conservative 41; Mismatches 71; Indels 48; Gaps 6;

QY 39 VPVTAVCLCLFVVGSGNVVTVMIGR---YRDMRTTNLYLGSMAVSDLLI-LLGLPFD 94
 DB 65 VLVTAYLALFVVGTVGNSVTAFTLARKKSLQSLQSTVHYHGLSLALSDDLILLAMPVE 124
 QY 95 LYSR-LWRSRPVWFGPGLLCRLSLYVGGCTVATLLHMTALSVERYLAICRPLRARVLVTR 153
 DB 125 LYNFIWVHHPWAGDAGCRGYFLRDCTATYATNALVASLSVERYLAICHFPKAKTLMRS 184
 QY 154 RVRLATVAVLWALLSAGPFLFLVGVQDPGISVVPGLNGTARIASSPPLMSRA 213
 DB 185 RTKKFISAIWLASALLAIPMLFTWGLQNRSGDGTGHPG-----GLVCTPIVDT----- 231
 QY 214 PPSPSPGPTAEAAALFSRECRPSPAQLGALRVMLVWTAYFEL-PFLCLSLYGLIGR 272
 DB 232 -----ATVKVVIQVNTFMSFLPMLVISILNTVIAN 262
 QY 273 EL 274
 DB 263 KL 264

RESULT 7
 NTRI MOUSE STANDARD; PRT; 424 AA.
 AC O88319; (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotensin receptor type 1 (NT-R-1).
 GN NTSR1 OR NTSR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Snider J., Sano H., Ohta M.;
 RT "Neurotensin receptor type 1";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR THE TRIPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

 CC EMBL: AB017027; BAA33013.1; --
 CC MGD; MGI:97386; Ncsl.
 CC DR InterPro; IPR000276; GPCR Rhodopsin.
 CC DR InterPro; IPR003985; NTL rec.
 CC DR InterPro; IPR003984; NTL rec.
 CC PFam: PF00001; Tm 1.1.
 CC DR PRINTS; PR00237; GPCRHOPOPSN.
 CC DR PRINTS; PR01479; NEUROTENSINR.
 CC DR PRINTS; PR01480; NEUROTENSINR.
 CC DR PROSITE; PS00237; G PROTEIN RECP F1.1; 1.
 CC DR PROSITE; PS0262; G PROTEIN RECP F1.2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 87 1 (POTENTIAL).
 FT DOMAIN 88 96 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 97 121 2 (POTENTIAL).
 FT DOMAIN 122 143 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 144 165 3 (POTENTIAL).
 FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 189 210 4 (POTENTIAL).
 FT DOMAIN 211 235 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 236 260 5 (POTENTIAL).
 FT DOMAIN 261 308 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 309 330 6 (POTENTIAL).
 FT DOMAIN 331 348 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 349 372 7 (POTENTIAL).
 FT DOMAIN 373 424 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 141 224 BY SIMILARITY.
 FT LIPID 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA; 47216 MW; 8E9A723171A48711 CRC64;

Query Match 16.8%; Score 340; DB 1; Length 424;
 Best Local Similarity 34.7%; Pred. No. 1.7e-16;
 Matches 84; Conservative 39; Mismatches 71; Indels 48; Gaps 7;

QY 39 VPVTAVCLCLFVVGSGNVVTVMIGR---YRDMRTTNLYLGSMAVSDLLI-LLGLPFD 94
 DB 64 VLVTAYLALFVVGTVGNSVTAFTLARKKSLQSLQSTVHYHGLSLALSDDLILLAMPVE 123
 QY 95 LYSR-LWRSRPVWFGPGLLCRLSLYVGGCTVATLLHMTALSVERYLAICRPLRARVLVTR 153
 DB 124 LYNFIWVHHPWAGDAGCRGYFLRDCTATYATNALVASLSVERYLAICHFPKAKTLMRS 183
 QY 154 RVRLATVAVLWALLSAGPFLFLVGVQDPGISVVPGLNGTARIASSPPLMSRA 213
 DB 184 RTKKFISAIWLASALLAIPMLFTWGLQ-----NRS 213
 QY 214 PPSPSPGPTAEAAALFSRECRPSPAQLGALRVMLVWTAYFEL-PFLCLSLYGLIGR 272
 DB 214 AUGQHPGG-----LVCTPT-VDTATVKVVIQVNTFMSFLPMLVISILNTVIAN 261
 QY 273 EL 274
 DB 262 KL 263

RESULT 8
 NTRI HUMAN STANDARD; PRT; 418 AA.
 ID NTRI1 HUMAN Q9H4T5;
 AC P30959; Q9H4H1; Q9H4T5;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucobastine-
 DE insensitive neurotensin receptor) (NTRH).

NTSR1 OR NTRR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93154505; PubMed=8381365;
 RA Vita N., Laurent P., Lefort S., Chalou P., Dumont X., Kaghad M.,
 RA Gully D., Le Fur G., Ferrara P., Caput D.;
 RT "Cloning and expression of a complementary DNA encoding a high
 RT affinity human neurotensin receptor.";
 RL FEBS Lett. 317:139-142(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehesvahti M.H., Leverhale M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McKay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulten J.E.,
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKINS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X70070; CAA49675.1; -;
 DR EMBL; AL357033; CAC14923.1; -;
 DR EMBL; AL035669; CAC12747.1; -;
 DR PIR; S29506; S29506.
 DR HSPG; P02699; IPF8.
 DR Gnew; HNCN:8039; NTSR1.
 DR MIM; 162651; -;
 DR GO; GO:0005783; C:endoplasmic reticulum; TAS.
 DR GO; GO:0005794; C:Golgi apparatus; TAS.
 DR GO; GO:0005987; C:integral to plasma membrane; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaling; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR InterPro; IPR00276; GPCR_Rhodopsin.
 DR InterPro; IPR003985; NTL1 Rec.
 DR InterPro; IPR003984; NTL1 Rec.

DR PFam; PF00001; 7tm1.1;
 DR PRINTS; PRO0237; GPCRHHODPSN.
 DR PRINTS; PRO1479; NEUROTENSINR.
 DR PRINTS; PRO1480; NEUROTENSINR.
 DR PROSITE; PS00327; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 63 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 64 86 1 (POTENTIAL).
 FT DOMAIN 87 95 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 96 120 2 (POTENTIAL).
 FT DOMAIN 121 142 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 143 164 3 (POTENTIAL).
 FT DOMAIN 165 187 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 188 209 4 (POTENTIAL).
 FT DOMAIN 210 234 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 235 259 5 (POTENTIAL).
 FT DOMAIN 260 303 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 304 325 6 (POTENTIAL).
 FT DOMAIN 326 343 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 344 367 7 (POTENTIAL).
 FT DOMAIN 368 418 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 141 224 BY SIMILARITY.
 FT LIPID 383 393 PALMITATE (POTENTIAL).
 FT CONFLICT 200 200 T -> A (IN REF. 2).
 SQ SEQUENCE 418 AA; 46288 MW; BBD1EBC2BE6E390 CRC64;
 Query Match 16.2%; Score 330; DB 1; Length 418;
 Best Local Similarity 31.2%; Pred. No. 8.3e-16;
 Matches 89; Conservative 44; Mismatches 88; Indels 64; Gaps 9;
 QY 39 VPVTAVCLFVVGSGNVVTVNLIGR---YRDMRTTNLYGSMVSDLL-ILLGLPPD 94
 DB 64 VLVTVAVLALFVVGTVGTVTAFTLARKKLSQSLQSTVHYHGLSLALSOLLTLLAMPVE 123
 QY 95 LYR-LWRSPWVPGPLLCRLSLVYVGCGTYATLLHMTALSVRYLAICRPLRARVLVTR 153
 DB 124 LYNFIVHHPWAFGDCRGYYFLRDACIYALNVASUSVERLYLAICHPFKAKTLMRS 183
 QY 154 RYRALIAVLWVALLSAGPFLVLVGVEQDPGIVSWPGLNGTARIASSPLLSRA 213
 DB 184 RTKFFISALWLASALITVPMFTMG-EQN----- 211
 QY 214 PPSPPSGPETAAALFSECRPSPAQGLARVLMVWTT-AVFPFLCLSLYLGLIGR 272
 DB 212 -----RSADQGHAGGLVCTPT-IHTATVKVVIQVNTFMSFIPMVVISLVNTIAN 261
 QY 273 ELWSSRRPL--RQPAAS-----GREGHRTQTVRLR 301
 DB 262 KLTVMYVQAQSQVCTVGEHSTFSMAIEPGRVQALRHGVRLR 306
 RESULT 9
 NTR2_HUMAN STANDARD; PRT; 410 AA.
 AC O95665; Q8TBH6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neurotensin receptor type 2 (NT-R-2) (Levocabastine-sensitive
 DE neurotensin receptor) (NTR2 receptor).
 GN NTSR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;

RX MEDLINE=99066919; PubMed=9851594;
 RA Vita N., Oury-Donat F., Chalton P., Guillemot M., Kaghad M., Bachy A.,
 RA Thurneysen S., Garcia S., Poinet-Chazel C., Casellas P., Keane P.,
 RA Le Fur G., Maffrand J.-P., Shoubrie P., Caput D., Ferrara P.,
 RT "Neurotensin is an antagonist of the human neurotensin NT2 receptor
 RL expressed in Chinese hamster ovary cells";
 RL Eur. J. Pharmacol. 360:265-272(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tohiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Kzyvinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y10148; CAA71233.1; -;
 DR EMBL; BC022501; AAH22501.1; -;
 DR Genbank; HGNC:8040; NTSR2.
 DR MIM; 605538; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0007600; P:sensory perception; TAS.
 DR InterPro; IPR000276; GPCR Rhodopsin.
 DR InterPro; IPR003986; NT2 rec.
 DR InterPro; IPR003984; NT rec.
 DR Pfam; PF00001; 7tm.1; 1;
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PRINTS; PR01479; NEUROTENSINR.
 DR PRINTS; PR01481; NEUROTENSINR.
 DR PROSITE; PS00237; G PROTEIN RECP FL 1; 1.
 DR PROSITE; PS00262; G PROTEIN RECP FL 2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 55 1 (POTENTIAL).
 FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 65 87 2 (POTENTIAL).
 FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 176 4 (POTENTIAL).
 FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 237 5 (POTENTIAL).
 FT DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 298 318 6 (POTENTIAL).
 FT DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 358 7 (POTENTIAL).
 FT DOMAIN 359 410 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 108 194 BY SIMILARITY.
 FT LIPID 377 377 PALMITATE (POTENTIAL).
 FT CONFLICT 168 168 L -> M (IN REF. 2).
 FT CONFLICT 367 367 K -> R (IN REF. 2).
 SQ SEQUENCE 410 AA; 45413 MW; 8C3ADA22BE15FD66 CRC64;
 Query Match 13.4%; Score 274.5; DB 1; Length 410;
 Best Local Similarity 32.6%; Pred. No. 5.2e-12;
 Matches 92; Conservative 36; Mismatches 91; Indels 63; Gaps 11;
 QY 17 PWPALPP---CDERCSFPFPGALVPTAVCLCLFVGVSGVNVVTVMLI-----GR 65
 DB 8 PRPSNPGSLDARIGVDVTRLWAKVLFALVALIWAAGNALSVVHVVKARAGR 67
 QY 66 YRDMRTTTLNLYGMAVSDLLIL-GLPFDLYR-LWRSRPWFVGLLRLSLVYVGGCTY 123
 DB 68 LRH-----HVLSLAGALLLVGVVELYSFVWFVFGDLGCRGYFVHELCA 120
 QY 124 ATLLHMTALSVRYLAICPLRVLVTRRRVRLIAVLMAVALLSAGPPLFLVGV--- 179
 DB 121 ATVLSVAGLSAERCLAVCCPLRARSLLTPRTRVLVLSWAASLGLALPMVIMGKH 180
 QY 180 -----EQDPSISVPGINGTARIAS-----PLASSPPL-----WLS 211
 DB 181 ETADGFEPSARVCTVL--VSRTALQVFIQNVNLSFVLPLALTAFNGVTVSHLLALCS 238
 QY 212 RAPPSPPPGPGPETAEEAALFSRECRPSAQLGALRVMLVWT 253
 DB 239 QVPSISTP-GSSTPSRLLELSEE-----GLLSFIVWKT 271
 RESULT 10
 NTR2 RAT
 ID NTR2 RAT STANDARD; PRT; 416 AA.
 AC Q63384;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotensin receptor type 2 (NT-R-2) (High-affinity leucobastine-
 DE sensitive neurotensin receptor).
 GN NTSR2 OR NTR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=96228041; PubMed=8647296;
 RA Chalton P., Vita N., Kaghad M., Guillemot M., Bonin J.,
 RA Delpech B., le Fur G., Ferrara P., Caput D.,
 RT "Molecular cloning of a leucobastine-sensitive neurotensin binding
 RT site";
 RL FEBS Lett. 386:91-94(1996).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER
 CC LEVELS SEEN IN THE HEART AND INTESTINE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
 CC EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
 CC BRAIN).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X97121; CAA65787.1; -
 FIR; S68822; S68822.
 InterPro: IPR000276; GPCR_Rhodopsin.
 InterPro: IPR003986; NT2_rec.
 InterPro: IPR003984; NT_rec.
 Pfam: PF00001; 7tm_1; 1_
 PRINTS; PR00237; GPCR_RHODOPSIN.
 PRINTS; PR01479; NEUROTENSIN.
 PRINTS; PR01481; NEUROTENSIN2.
 PROSITE; PS00237; G PROTEIN RECP Fl_1; 1.
 PROSITE; PS0262; G PROTEIN RECP Fl_2; 1.
 G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32
 FT TRANSMEM 33 55
 FT DOMAIN 56 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 176
 FT DOMAIN 177 216
 FT TRANSMEM 217 237
 FT DOMAIN 238 297
 FT TRANSMEM 298 318
 FT DOMAIN 319 337
 FT TRANSMEM 338 358
 FT DOMAIN 359 416
 FT DISULFID 108 194
 FT LIPID 377 377
 SEQUENCE 416 AA; 46265 MW; 127C5F5CB8FE208 CRC64;

Query Match 13.1%; Score 268.5; DB 1; Length 416;
 Best Local Similarity 37.9%; Pred. No. 1.4e-11;
 Matches 66; Conservative 33; Mismatches 66; Indels 9; Gaps 4;
 QY 18 PWALP-----PCDERRCSPFPGALVPTAVCLCLFVGVGSGNVVTVMLGRYDRMT 71
 DB 6 PWPFRSPSAGLSLEARLVDTLWAKVLTALYSLIFAFGTAGNALSVHVLKARGRP 65
 QY 72 -TTNLVGSMAVSDLLILL-GLPFDLYR-LWRGRPWVFGPLLCRLSLYVGEGCTATLLH 128
 DB 66 GLRYHVLSLALGALLLVSMPELYNFWVSHYPVFGDLGCRGYFVRELCAVATLS 125
 QY 129 MTALSVRYLAICRLRRLVTRRRVRLALVAVALLSAGPFLFLVGVEQD 182
 DB 126 VASLSAERCLAVCQPLRARELLTPRTTRRLSLVWVASLGLALPMAVIMQKHE 179

RESULT 11
 GSTR_BOVIN STANDARD; PRT; 454 AA.
 AC P79266;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Gastrin/cholecystokinin type B receptor (CKC-B receptor) (CKC-BR).
 GN CCKBR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=97003869; PubMed=8851180;
 RA Dufréne M., Escrieu C., Clerc P., le Huerou-Luron I., Prats H.,
 Bertrand V., le Meuth V., Guillotheau P., Vaysse N., Fourmy D.,

"Molecular cloning, developmental expression and pharmacological characterization of the CCKB/gastrin receptor in the calf pancreas.", Eur. J. Pharmacol. 297:165-179(1996).
 -1- FUNCTION: RECEPTOR FOR GASTRIN AND CHOLECYSTOKININ. THE CCK-B RECEPTORS OCCUR THROUGHOUT THE CENTRAL NERVOUS SYSTEM WHERE THEY MODULATE ANXIETY, ANALGESIA, AROUSAL, AND NEUROLEPTIC ACTIVITY. THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 -1- SUBCELLULAR LOCATION: Integral membrane protein.
 -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S83030; AAB46896.1; -
 HSPB; P02699; 1F88.
 InterPro: IPR000276; GPCR_Rhodopsin.
 Pfam: PF00001; 7tm_1; 1_
 PRINTS; PR00237; GPCR_RHODOPSIN.
 PROSITE; PS00237; G PROTEIN RECP Fl_1; 1.
 PROSITE; PS0262; G PROTEIN RECP Fl_2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein;
 Lipoprotein; Palmitate; Phosphorylation.
 FT DOMAIN 1 57
 FT TRANSMEM 58 79
 FT DOMAIN 80 87
 FT TRANSMEM 88 109
 FT DOMAIN 110 131
 FT TRANSMEM 132 150
 FT DOMAIN 151 170
 FT TRANSMEM 171 189
 FT DOMAIN 190 219
 FT TRANSMEM 220 242
 FT DOMAIN 243 340
 FT TRANSMEM 341 362
 FT DOMAIN 363 380
 FT TRANSMEM 381 401
 FT DOMAIN 402 454
 FT CARBOHYD 7 7
 FT CARBOHYD 30 30
 FT CARBOHYD 36 36
 FT DISULFID 127 205
 FT LIPID 415 415
 SEQUENCE 454 AA; 48781 MW; A2846A580508ABA6 CRC64;

Query Match 13.1%; Score 268; DB 1; Length 454;
 Best Local Similarity 25.6%; Pred. No. 1.6e-11;
 Matches 99; Conservative 60; Mismatches 139; Indels 88; Gaps 14;
 QY 2 GSP-WNGSDGPEGAREPPWPPALPCDERRCSPFPGALVPTAVCLCLFVGVGSGNVTV 60
 DB 25 GGPLNGSGTGNLSCEPP-----RIRGAGTRELALRVLVAV-IFLMSVGVNVLII 76
 QY 61 MLIGRYDRVTRTNLYLGSMAVSDLLILIG-LPFDLYRLWRGRP-----WVFGPLLCR-L 113
 DB 77 VVLGSRRLRTVTNAPFLSLAVSDLLAVACMPFTLL-----PNLMGTPIFGTVVCKAV 130
 QY 114 SLVVGEGCTATLLHMTALSVRYLAICRLRRLVTRRRVRLALVAVALLSAGPFL 173
 DB 131 SYFMGVSVSVST-LSLVAIALERYSAICRPLQARVQWTRSHAARVIVATMTLSGLLV 199
 QY 174 LFLVGVQDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPSPSPGPGTAAALFSR 233
 DB 190 PVTAVQ-----PAGPRVLOCMRW-- 209
 QY 234 ECRPSPAQLGALRVLMLVTVTAYFFLPFLCLSLYGLIGELM-----SSRRP 280

Db 210 ---PSARVRQTSVLLILL--LLFFVPGVMNAVGLISRELYLGLRFDGSDSESQSRVG 264

Qy 281 LRGPAAAGRGHRQTVRLKWSRGSK--DACLOSAPGTAQTGLPLLLAQLWAPLP 338

Db 265 SOGLPGGTGGPAQANGRCSETRIAGEDGCGYQLPRSR-----PALEMSALTATPT 319

Qy 339 APPISIPASTRRGGGGIYNLLVAL 364

Db 320 GPSGGRPAQAKLLAKKRVVRLVI 345

RESULT 12

OPRX_CAVPO STANDARD; PRT; 370 AA.

AC P47748; 1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Nociceptin receptor (Orphanin FQ receptor) (Kappa-type 3 opioid receptor) (KOR-3) (XOR).

GN OPR1 OR OOR

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hartley; TISSUE=Brain;

RA Xie G.;

RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ. HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS, INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLIL CYCLASE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U04369; AAA58332.1; -

DR HSSP; P34996; 1DD0.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm.1.1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.

DR PROSITE; PS0262; G-PROTEIN RECP F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 51 77 1 (POTENTIAL).

FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 88 109 2 (POTENTIAL).

FT DOMAIN 110 124 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 125 146 3 (POTENTIAL).

FT DOMAIN 147 165 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 166 188 4 (POTENTIAL).

FT DOMAIN 189 211 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 212 236 5 (POTENTIAL).

FT DOMAIN 237 264 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 265 288 6 (POTENTIAL).

FT DOMAIN 289 300 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 301 322 7 (POTENTIAL).

FT DOMAIN 323 370 CYTOPLASMIC (POTENTIAL).

FT DISULFID 123 200 BY SIMILARITY.

FT LIPID 334 334 PALMITATE (POTENTIAL).

FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 370 AA; 40789 MW; DAA807CE24283573 CRC64;

Query Match 13.1%; Score 267.5; DB 1; Length 370;

Best Local Similarity 24.9%; Pred. No. 1.4e-11;

Matches 83; Conservative 53; Mismatches 116; Indels 81; Gaps 7;

Qy 17 PMPALPP---CDERRCSPPFLGALVPVTAVCGLCFVWVGSGNVVTVMLIGRYDRMTT 73

Db 27 PNHSGLPFPHLLLNASHAFPLGLKVTIVGLYLAVCIGGLGNCCLVMYVILRHTKMTAT 86

Qy 74 NLYLGSMAVSDLLILGLPFDLYRLWRSRPVFGPLLCRLSLYVVGSGCTVATLLHMTALS 133

Db 87 NIYIFNLALADTLVLLTLPFOATDILLGF-WFPGNTLCKTVTAIDYNNFTSTFTTAMS 145

Qy 134 VERYLAICRPLRARVLVTRRRVRLAVLWAVALLSAGPFLFLVGVQDPGSGVWPLNG 193

Db 146 VDRYVAICHPIRALDVRTSSKAQAVNAIWAALV-----VGVPVAIMG 189

Qy 194 TARIASSPLASSPPLWLSRAPPPSPSPSPGPETAEAAALFSRECPSPAQLGALRVMLWVT 253

Db 190 SAQVEDDEIEC-----LVRIPIPDQDYMGVFAVSIPLFS----- 223

Qy 254 AYFFLPFLCLSLYGLIGRELWSSR-----RPLRG----- 283

Db 224 --FIPLVLIISVCYSLMIRRLHGVRLLSGSRKDRNLRLVLLVVAVFVGCWTPVQV 281

Qy 284 ---PAASGRGHRQTVRLKWSRGSKDACL 313

Db 282 FVLVQGLGVQPGSETTVALRFTCALGYVNSCL 314

RESULT 13

NTR2_MOUSE STANDARD; PRT; 417 AA.

AC P70310;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Neurotensin receptor type 2 (NT-R-2) (Low-affinity leucobastine-sensitive neurotensin receptor) (NTRL).

GN NTR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Brain;

RX MEDLINE=96388216; PubMed=8795617;

RA Mazella J., Botto J.-W., Guillemaire E., Coppola T., Sarret P., Vincent J.-P.;

RA "Structure, functional expression, and cerebral localization of the leucobastine-sensitive neurotensin/neuromedin N receptor from mouse brain."

RT J. Neurosci. 16:5613-5620(1996).

RL -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED MAXIMALLY IN THE CEREBELLUM, HYPOCAMPUS, PIRIFORM CORTEX AND NEOCORTEX OF ADULT BRAIN.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED POORLY IN 7-DAY-OLD BRAIN. EXPRESSION INCREASES AT DAY 15 TO REACH A MAXIMAL LEVEL IN 35-DAY-OLD BRAIN.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININS RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; U51908; AAB17285.1; -
CC MGD; MGI:108018; Nter2
CC InterPro; IPR000276; GPCR_Rhodopsin
CC InterPro; IPR003986; NT2_rec
CC InterPro; IPR003984; NT_rec
CC Pfam; PF00001; 7tm.1; 1
CC PRINTS; PR00237; GPCR_Rhodopsin
CC PRINTS; PR01479; NEUROSENSIN
CC PRINTS; PR01481; NEUROSENSIN2
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1
CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1
CC G-protein coupled receptor; Transmembrane; Palmitate.
FT DOMAIN 1 32
FT TRANSMEM 33 55
FT DOMAIN 56 64
FT TRANSMEM 65 87
FT DOMAIN 88 109
FT TRANSMEM 110 131
FT DOMAIN 132 154
FT TRANSMEM 155 176
FT DOMAIN 177 217
FT TRANSMEM 218 238
FT DOMAIN 239 298
FT TRANSMEM 299 319
FT DOMAIN 320 338
FT TRANSMEM 339 359
FT DOMAIN 360 417
FT DISULFID 108 194
FT LIPID 378 378
SQ SEQUENCE 417 AA; 46537 MW; BFFDDBD650723DD CRC64;

Query Match 13.0%; Score 266; DB 1; Length 417;
Best Local Similarity 32.5%; Pred. No. 2e-11;
Matches 81; Conservative 36; Mismatches 88; Indels 44; Gaps 8;

QY 19 WPALP-----PCDERCSPPFGLALVPTAVCLFVGVSGNVVTVMLIGRYVDMRT- 71
Db 7 WPPRPSAGLSLEARGVDTLWAKVLFYALYSLIFALGTAGNLSVHVVLKARTGPG 66
QY 72 TTNLYLGSMAVDLLILL-GLPFDLYR-LMRSEFWFGLLCRLSLYVGGCTVATLLHM 129
Db 67 RLRYHLSLALSALLLSVPMELYNFVWSHYPWVFDLGCGRYFVRELCAVATVLSV 126
QY 130 TALSVERYLAICPLRARVLVTRRRVATLAVLWAVALLSAGPFLVGVQDPGIVSWP 189
Db 127 ASLSAERCLAVCPQLRARLLTPRTCTRLSLVWVASLGLALPMVIMGQKHE----- 179
QY 190 GLNGTARIASSPLASPPMLSRAP-PPSPGSPETAEEAALFSRECRPSPAQL---GAL 245
Db 180 -----MERADGPEPAS-----RVCTVLVSRASSRSTFQVKRAGLL 215
QY 246 RVMWLWTTA 254
Db 216 RSLPWLTA 224

RESULT 14
ID SSR5 MOUSE STANDARD; PRT; 362 AA.
AC O08850; O08998;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Somatostatin receptor type 5 (SSR5).
GN SSR5 OR SMSTR5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj; TISSUE=Liver;
RX MEDLINE=97444289; PubMed=9300821;
RA Lublin A.L., Diehl N.L., Hochgeschwender U.;
RT "Isolation and characterization of the gene encoding the type 5 mouse
RL (Mus musculus) somatostatin receptor (msst5).";
RN Gene 195:63-66 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj;
RX Moldovan S., Demayo F., Brunicardi P.C.;
RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX Gordon D.F., Woodhouse W.W., Wood W.M., Knauf H., James R.A.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj;
RX Baumeister H., Roosterman D., Schafer J., Kreuzer O., Meyerhof W.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Query Match

12.8%; Score 262; DB 1; Length 362;

Best Local Similarity 26.1%; Pred. No. 3.3e-11;
Matches 80; Conservative 45; Mismatches 103; Indels 78; Gaps 10;

QY 5 WNGSDGEGAREPPWALPCDERRCSPPLGA---LVPYAVCLCLFVVGSGNVVTVM 61
DB 12 WNASASSGSHN--WSLVDPVS-----PMGARVLVPVLYLVC--TVGLGGLTVIY 60

QY 62 LIGRYDMRTTNNLYGSMVSDLLILLGLPFDLYRLWRSPWVFGPFLCRLSLYVGE 121
DB 61 VVLYAKMKTIVYILNLAVADVFLMLGLFF-LATQNAVSYWPFSGFLCRLVMTLDGIN 119

QY 122 TYATLHMTALSVERYLAICRPLARVLVTRRRVRAIAVLMAVALISAGPFLFLV 181
DB 120 QFTSIFCLMWSVDYLVAVHPLRSARWRPRVAKLASAAVMVFSLLMSPLLVFAD 179

QY 182 DPGISVVGLNGTARTIASSPLASPPPLWLSRAPPPSPGPETAETAAALFSR 241
DB 180 G-----WGTCNLS-----W-----PEPVG 193

QY 242 LGALRVMLVTTTAYFFLPFLCLSLYGLI-----GRELWSSRRPLRGPASGR 294
DB 194 LWGAFTYTSVLGFFGLVLCVLLIVVUKAAGRVGSSR-----RRRSR 243

QY 295 QTVRVL 300
DB 244 KVTMRV 249

RESULT 15
OPRX_PIG STANDARD; PRT; 370 AA.

AC P79292;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nociceptin receptor (Orphanin FQ receptor) (Kappa-type 3 opioid
DE receptor) (KOR-3) (ORGC) (K3 opiate receptor).
GN ORPL1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RA MEDLINE=99140612; PubMed=9988113;
RX Osinski M.A., Pampusch M.S., Murtaugh M.P., Brown D.R.;
RT "Cloning, expression and functional role of a nociceptin/orphanin FQ
RT receptor in the porcine gastrointestinal tract.";
RL Eur. J. Pharmacol. 365:281-289(1999).
CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYL
CC CYCLASE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; U72758; AAB39702.1; -;
CC HSSP; P34996; 1DDD.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.

G-protein coupled receptor; Transmembrane; Glycoprotein;
Phosphorylation; Lipoprotein; Palmitate.

KW DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 77 1 (POTENTIAL).
FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 146 3 (POTENTIAL).
FT DOMAIN 147 165 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 166 188 4 (POTENTIAL).
FT DOMAIN 189 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 236 5 (POTENTIAL).
FT DOMAIN 237 264 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 265 288 6 (POTENTIAL).
FT DOMAIN 289 300 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 301 322 7 (POTENTIAL).
FT DOMAIN 323 370 CYTOPLASMIC (POTENTIAL).
FT DISULFID 123 200 BY SIMILARITY.
FT LIPID 334 334 PALMITATE (POTENTIAL).
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 370 AA; 48E4F3B6DA223BCD CRC64;

Query Match 12.5%; Score 254.5; DB 1; Length 370;
Best Local Similarity 26.1%; Pred. No. 1.1e-10;
Matches 80; Conservative 49; Mismatches 114; Indels 63; Gaps 8;

QY 2 GSPWNGSDGEGAREPPWALPP---CDERRCSPPLGALVPVTAVCLCLFVVGSGNVV 58
DB 15 GSPLOQN---LSLLSPNHSLLPPLLLNASHGAFPLGLKVTIVGLYLAVCVGGLGNC 71

QY 59 TVMLIGRYDMRTTNNLYGSMVSDLLILLGLPF---DLYRLWRSPWVFGPFLCRLSL 115
DB 72 VMXVILRHTKMTATNIYIFNLALADTAVLLTLPFGQTDVLGF-----WPFGNALCKAVI 127

QY 116 YVGECTYATLHMTALSVERYLAICRPLARVLVTRRRVRAIAVLMAVALISAGPFLF 175
DB 128 AIDYNNMFTSAFTLTAMSDRYVAICHPIRALDVRTSSKAQAVNVAIWALASI----- 180

QY 176 LVGVEODPGISVVGLNGTARTIASSPLASPPPLWLSRAPPPSPGPETAETAAALFSR 235
DB 181 -----VGVPVAIMGSAQVEDEIEC-----LVEIPAPQDYMGPVFAVCIFLFS--- 223

QY 236 RSPQAQLGALRVMLVTTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPASGRGRHQ 295
DB 224 -----FVIPVLIISVCYSLMVRRLRGVR-----LLSGSRKORN 257

QY 296 TVRVLR 301
DB 258 LRRIIR 263

Search completed: January 1, 2004, 06:23:02
Job time : 14.5439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 63.3659 Seconds
(without alignments)
1571.954 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDGPEGAREPWP.....WQNLHKGKGFADVLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	673.5	33.0	363	13	O93413 spherooides
2	641.5	31.4	374	13	O93412 spherooides
3	641	31.4	364	11	Q8BWX8 mus musculus
4	607	29.7	366	6	Q8MHZ5 ovis aries
5	435	21.3	559	13	O93414 spherooides
6	356.5	17.4	426	4	Q8NE20 homo sapien
7	354.5	17.4	426	4	Q9HB89 homo sapien
8	339	16.6	145	6	Q97914 ovis aries
9	338	16.5	395	11	Q91276 mus musculus
10	338	16.5	395	11	Q8BZ39 mus musculus
11	336	16.4	415	4	Q96AM5 homo sapien
12	335	16.4	415	4	Q9GZ04 homo sapien
13	331.5	16.2	405	11	O55040 mus musculus
14	329	16.1	395	11	Q9JIB1 rattus norv
15	329	16.1	395	11	Q9ESQ4 rattus norv
16	316	15.5	477	5	Q8ITC7 drosophila

17	309	15.1	428	5	Q9VFN4	Q9vfn4 drosophila
18	309	15.1	430	5	Q8ITC9	Q8itc9 drosophila
19	304.5	14.9	412	11	Q9JJI5	Q9jji5 rattus norv
20	302.5	14.8	402	11	Q9JIB2	Q9jib2 rattus norv
21	285.5	14.0	418	5	O17239	O17239 caenorhabdi
22	266.5	13.0	401	13	Q8JF26	Q8jif26 xenopus lae
23	264.5	12.9	401	13	Q9DDR0	Q9ddr0 xenopus lae
24	262.5	12.8	416	11	Q8VIFS	Q8vifs mus musculus
25	262.5	12.8	416	11	Q920Q5	Q920q5 mus musculus
26	262	12.8	385	11	Q9JK40	Q9jk40 mus musculus
27	261	12.8	595	5	Q8ITD0	Q8itd0 drosophila
28	261	12.8	599	5	Q9VFW6	Q9vfw6 drosophila
29	255.5	12.5	378	5	Q18701	Q18701 caenorhabdi
30	253.5	12.4	516	4	Q96LC6	Q96lc6 homo sapien
31	252.5	12.4	660	5	Q9VFW5	Q9vfw5 drosophila
32	251.5	12.3	516	4	Q9NVK7	Q9nvk7 homo sapien
33	250.5	12.3	513	13	Q9DPA9	Q9dpa9 cacoctomus
34	250.5	12.3	658	5	Q8ITD1	Q8itd1 drosophila
35	250	12.2	477	13	Q8JID5	Q8jid5 carassius a
36	249	12.2	367	6	Q8MI04	Q8mi04 ovis aries
37	248	12.1	453	11	Q8BKF6	Q8bkf6 mus musculus
38	246.5	12.1	314	11	Q8BYG7	Q8byg7 mus musculus
39	245.5	12.0	382	11	Q9ERT2	Q9ert2 mus musculus
40	241	11.8	370	13	Q8UWLS	Q8uwl5 fugu rubrip
41	240.5	11.8	371	11	Q8BKB0	Q8bkb0 mus musculus
42	240	11.7	352	6	Q8HZP1	Q8hzp1 tupaia mino
43	240	11.7	450	11	P89005	P89005 praomys nat
44	239.5	11.7	517	5	Q8TOS8	Q8tos8 drosophila
45	239	11.7	352	11	O88820	O88820 rattus norv

ALIGNMENTS

RESULT 1

O93413 PRELIMINARY; PRT; 363 AA.
AC O93413;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP MEDLINE=20092336; PubMed=10628755;
RA Palyha O.C., Feigener S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
secretagogue receptor (GHS-R) conserved from pufferfish to humans."
RL Mol. Endocrinol. 14:160-169(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF082210; AAC33473.1; -;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 363 AA; 41324 MW; 74518BD9CD310991 CRC64;

Query Match 33.0%; Score 673.5; DB 13; Length 363;

Best Local Similarity 49.8%; Pred. No. 5.9e-45;

Matches 135; Conservative 39; Mismatches 55; Indels 43; Gaps 4;

QY 31 SPPLGALVPVTAVCICLFVVGSGNVVTVMLIGRYDRMRTTNNLYLGSMVSDLLILG 90

Db 31 SLFPASTLIPVTICILIFVVGVTGNTMTIIQYFKDMKTNTNLVSSMAVSDLVFLC 90
 QY 91 LPDLYLWLSRPWVFGCLRLSLVYEGCTATLHMTALSVERVLAICRPLRARVLY 150
 Db 91 LPDLYLWLVKVPWFGAEVCRVHYIFEGCTATLHMTALSVERVLAISFFPLRSKVMV 150
 QY 151 TRRRVRLIALVMAVALLSAGPFLFLVGVQDPGISVVVGLNGTARIASSPLASSPPLWL 210
 Db 151 TRRRVQVILALMCFALVSAAPTILFVGVYD-----NET----- 185
 QY 211 SRAPPPSPGPPETAALFSSRECRSPAQL--GALRVMLVWTTAYFFLPFLCLSLIYG 268
 Db 186 ----HPDNTG-----QCKHTGYAISGQLHIMVSTVYFFCPMLCLLFLYG 229
 QY 269 LIGRELWSSRPLRGPAAASRGHROTQVRL 300
 Db 230 SIGCKLWKSNDLQGPCALARSHROTQVIL 261

RESULT 2

ID Q93412 PRELIMINARY; PRT; 374 AA.
 AC Q93412;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Orphan G protein-coupled receptor.
 OS Spheroidea nepelus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Spheroidea.
 OX NCBI_TaxID=39110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2092336; PubMed=10628755;
 RA Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
 RA Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
 RA Patchett A.A., Howard A.D., Smith R.G.,
 RT "Ligand activation domain of human orphan growth hormone (GH)
 RT secretagogue receptor (GHS-R) conserved from pufferfish to humans."
 RL Mol. Endocrinol. 14:160-169(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AF082209; AAC33472.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR00524; HTH_GntR.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G PROTEIN RECP FL 1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECP FL 2; 1.
 DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 374 AA; 42324 MW; 2CF9304F004C7A16 CRC64;

Query Match 31.4%; Score 641.5; DB 13; Length 374;
 Best Local Similarity 45.8%; Pred. No. 2e-42;
 Matches 138; Conservative 45; Mismatches 73; Indels 45; Gaps 8;
 QY 5 WNGS-DGPEGARPPPALPCDERRCSPFPLGALVPVAVCLCLFVVGSGNVVTVMLI 63
 Db 13 WEGSHNGTAGLEUPP-----LNYSIPLAVITVACTVLTGVGVGNVTVILVV 61
 QY 64 GRYRDMRTTNLYLGSNAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSLYVGGCTY 123
 Db 62 SYRDMRTTNLYLGSNAVSDLFTVCMPLDLYRMWRYPWRFGDALCKLFQVSESTY 121
 QY 124 ATLLHMTALSVERVLAICRPLRARVLTTRRRVRLAVLWALLSAGPFLFLVGVQDP 193
 Db 122 STILICLALSVERVLAICFPRLAKVTKRRVRLAILLLTVSLLSAGPFFVVMVGEKDS 191
 QY 184 GISVFGNGTARIASSPL-ASSPPLWLSRAPPPSPGPTAEAAALFSRECRPS--PA 240

Db 182 --IMFPN-----SSDLNESSWPL-----BAVDTRCRMTOYAV 212
 QY 241 QIGALRVMLVWTTAYFFLPFLCLSLIYGLIGRELWSSRPLR-GPAASRGHROTQVRY 299
 Db 213 ESLGMEAMVWLSVFFMPVFCVLTVLYGLIGRRLWLRHRETTINSRVAYRDKSNROTIRM 272
 QY 300 L 300
 Db 273 L 273
 RESULT 3
 ID Q8BW8 PRELIMINARY; PRT; 364 AA.
 AC Q8BW8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Growth hormone secretagogue receptor type 1 homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK049671; BAC33866.1;
 SQ SEQUENCE 364 AA; 40969 MW; 8F1214B58F3B2E8 CRC64;

Query Match 31.4%; Score 641; DB 11; Length 364;
 Best Local Similarity 45.8%; Pred. No. 2.1e-42;
 Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;
 QY 5 WNGSDGPEGAREP-----PWPALPCD---RRRCSPFPLGALVPVAVCLCLFVVGSG 55
 Db 2 WNAT--PSEPEPNVTLDLDWDASPGNDSDELLPFPAPLAGVATATCVALFVVGISG 59
 QY 56 NVVTWMLICRYDMRTTNLYLGSNAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSL 115
 Db 60 NLLTMLVRSFRRLRTTNLYLSSMAFSDLLILFCLMPLDLVRLWQYRPNFGLLCKLFQ 119
 QY 116 YVGEGETATLLHMTALSVERVLAICRPLRARVLTTRRRVRLAVLWALLSAGPFLF 175
 Db 120 FVSECTYATVLTITALSVERVLAICFPRLAKVTVTKGRVILVILWAVAFCSAGPIFV 179
 QY 176 LVGVEODPGISVVVGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSREC 235
 Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
 QY 236 R88--PAQIGALRVMLVWTTAYFFLPFLCLSLIYGLIGRELWSSRPLRCPAASG---RE 290
 Db 198 RATEPAVRSGLLTVMWVSSVFFFLPVFCLTVLYSLIGRKLW--RR--RGDAVAGSSLRD 253
 QY 291 RGHROTQVRL 300
 Db 254 QNHKQIVKML 263
 RESULT 4
 ID Q8MHZ5 PRELIMINARY; PRT; 366 AA.
 AC Q8MHZ5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Ghrelin/growth hormone secretagogue receptor.

```
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=1, 2, and 3; TISSUE=Pituitary;
RA Murata T., Ikegami R., Morita Y., Shinozaki K.;
RT "Sheep Ghrelin/Growth Hormone Secretagogue Receptor.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY093948; AAM19733.1; -
DR EMBL; AY093949; AAM19734.1; -
DR EMBL; AY093950; AAM19735.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR004704; IID_comp.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF03613; EIID-AGA; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECP FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 366 AA; 41486 MW; 2F276651BC6C5D57 CRC64;

Query Match 29.7%; Score 607; DB 6; Length 366;
Best Local Similarity 42.2%; Pred. No. 9.6e-40;
Matches 129; Conservative 46; Mismatches 79; Indels 52; Gaps 6;

QY 5 WNCSDGPE-----GAREPPWALPPCD---ERRCSPPGALVPVTAICLFLVVGSGN 56
DB 2 WNA7SEELGNLTLPDLDDAAPDNDLSLTDELPPFPAPLAGVATATCALVFWGAGN 61

QY 57 VVTVMILIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRWVFGPLLRLSLY 116
DB 62 LITLVVSRFREERTTNLYLSWAFSDLLIFLCPDLVRLWHYRPWNLGDLCLKLQF 121

QY 117 VEGCTYATLLHMTALSVERLYLAICRLARVLVTRRRVRLIAVLMAVALLSAGPFLFL 176
DB 122 VSESCTVASVLTITALSVERLYFAICPLRAKVVITKGRVGLVLAIVAVAFCSAWPIFML 181

QY 177 VGVQDPCISVVPGLNGTARIASSPLASSPPLMSRAPPSGPTAEAAALFREC 236
DB 182 VGVGHE-----NGT-----DPRD-----TNECR 199

QY 237 PS--PAQALGALVWLWTTAYFFLPFLCLSLYGLIGLWSSRRPLRGPAAASRGHR 294
DB 200 ATFAVRSGLLTIMVWSSIFFLPVFLVCLYSLIGRLWRRRRSEVVVGASLRDQNHK 259

QY 295 QTVRVL 300
DB 260 QTVKML 265

RESULT 5
O93414 PRELIMINARY; PRT; 559 AA.
AC O93414;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20092336; PubMed=10628755;
RA Palya O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schlem K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
```

```
RT "Ligand activation domain of human orphan growth hormone (GH)
secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
DR EMBL; AF082211; AAC33474.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECP FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 559 AA; 61956 MW; AFB9F35FAF0CD5F6 CRC64;

Query Match 21.3%; Score 435; DB 13; Length 559;
Best Local Similarity 31.2%; Pred. No. 4.5e-26;
Matches 111; Conservative 52; Mismatches 109; Indels 84; Gaps 9;

QY 22 LPFCDERRC---SPPLGALVPVTAICLFLVVGSGNVVTVMLIGRYDMRTTNLYLG 78
DB 20 LHKCSNQECHWEEPV-FGMIVCVTIYIPLMFLGLGNLITLVVWLRYMRSTLYLIS 78

QY 79 SMAVSDLLILGLPDLVRLWRSRWVFGPLLRLSLYVGECTYATLLHMTALSVERLY 138
DB 79 SLAVSDILILLLLPVDLYLWPRWPFGEIECKSTMFSECTFCISILHITFISLERYL 138

QY 139 AICRLARVLVTRRRVRLIAVLMAVALLSAGPFLVGVG-----QDPGIS----- 186
DB 139 AVCWPTAKTVTRRTTRTIIGCIWLGAALSAAPVWVVGVEVQSDQGLSGWRESGAW 198

QY 187 -----VBPGL----- 191

DB 199 TKGEKGFIIGERERENDKGLKDLQLEMMNWKEMNECGDKNGVTGFKGGDKSLEV 258

QY 192 -NGTARIASSPLASSPPLMSRAPPSGPE-----TAEEAALFSGRCRPS--PAQLGA 244
DB 259 GEGTKEQEHGEAAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE 318

QY 245 LRVLWVTTAYFPFLPFLCLSLYGLIGLWSSRRPLRGPAAASRGHRQTVRVL 300
DB 319 LSAMLVLSNMVFLVPLCILGLVGLIGRLTW-----LRS-QISRRDVRNRTVKML 368

RESULT 6
Q8NE20 PRELIMINARY; PRT; 426 AA.
ID Q8NE20;
AC Q8NE20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G protein-coupled receptor 66.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC036543; AAH36543.1; -
DR Genex; HGNC:4518; GPR66.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR005390; NeuromedinU.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01566; NEUROMEDINUR.
DR PROSITE; PS00237; G PROTEIN RECP FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 426 AA; 47322 MW; 7BFEB5DE2E2DC686 CRC64;

Query Match 17.4%; Score 356.5; DB 4; Length 426;
```


Qy	287	SGRGRGH	293	
		:		
Db	139	SLRDQNH	145	
RESULT 9				
Qy	Q91Z76	PRELIMINARY;	PRT;	395 AA.
ID	Q91Z76			
AC	Q91Z76;			
DT	01-DEC-2001	(TrEMBLrel. 19, Created)		
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	Neuromedin U receptor type 2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_TaxID	10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RA	Funes S., Hedrick J.A., Yang S., Shan S., Bayne M., Monema F.J. Jr.,			
RA	Gustafson E.L.;			
RT	"Characterization of murine neuromedin U R2 receptor.";			
RL	Submitted (SP-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
DR	EMBL; AY057384; AAL26695.1; -			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	InterPro; IPR005390; NeuromedinU.			
DR	InterPro; IPR005392; NeuromedinU2.			
DR	Pfam; PF00001; 7tm1.1.			
DR	PRINTS; PR00237; GPCRRHODOPSN.			
DR	PRINTS; PR01565; NEUROMEDINUR.			
DR	PRINTS; PR01567; NEUROMEDINUR.			
DR	PROSITE; PS00237; G-PROTEIN RECEPT_F1.1; 1.			
DR	PROSITE; PS00262; G-PROTEIN RECEPT_F1.2; 1.			
KW	G-protein coupled receptor; Receptor; Transmembrane.			
Qy	SEQUENCE	395 AA; 44844 MW; 0BB540024566903D CRC64;		
Query Match 16.5%; Score 338; DB 11; Length 395;				
Best Local Similarity 30.9%; Pred. No. 1.3e-18;				
Matches 77; Conservative 54; Mismatches 72; Indels 46; Gaps 5				
Qy	30	CSPPLGLALVPVAVCLCLFVVGVSQNVVWMLICRYRDMTNTNLYLGSAVSDLLI-L	88	
Db	29	CGPKRSDLSLPVSVVYALIFVVGIVGNLLVCLVIARHQTLPKTPNTNYLFLSLAVSDLLVLL	88	
Qy	89	LGLPFDLYLRWSRPWFPGPLLCRLSLVVGECTYATLLHMTALSVERYLAICRPLRARV	148	
Db	89	LGMPLVVELHNNYFPFLGPGVCYFKTALFETVCFASILSVTTVSIERVVAIVHPFRAKL	148	
Qy	149	LVTRRRVVALIAVLWAVALLSAGPFLFVG--EQDPGISVVVPGNGTARTIASSPLASP	206	
Db	149	ESTRRRLRLISLWMSFSVWFSLPNTSHIGIKFQQFPNGSSVPG-----SATCTVTK	200	
Qy	207	PLWLSRAPPPSPGPEATAEAALFSRECRSPAQLGALRVLWLTATYFF-LPFLCLSI	265	
Db	201	PMWV-----YNFIQATSEFLFPLPMTLISV	226	
Qy	266	LYGLIGREL	274	
Db	227	LYLMGLRL	235	
RESULT 10				
Qy	Q9BZ39	PRELIMINARY;	PRT;	395 AA.
ID	Q9BZ39			
AC	Q9BZ39;			
DT	01-MAR-2003	(TrEMBLrel. 23, Created)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	Neuromedin U receptor type 2.			

DR InterPro; IPR005392; NeuromedinU2.
DR Pfam; PF00001; 7cm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS01565; NEUROMEDINUR.
DR PROSITE; PS01567; NEUROMEDINU2.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;

Query Match 16.1%; Score 329; DB 11; Length 395;
Best Local Similarity 30.1%; Pred.No. 6.4e-18;
Matches 75; Conservative 55; Mismatches 73; Indels 46; Gaps 5;

Qy 30 CSPFPPLGALVPVAVCLCFPVGVSGNVVTMLIGRYRDMRTTNNLYLGSMAVSLLLI-L 88
Db 29 CGPKRSDLSPVSVAYALIFLVGMGNLLVCWIVRHQTLPKTPNYLPSLAVSDLLVL 88
Qy 89 LGLPFDLRLWRGRPMVGPPLLCRLSLYVGEGCTYATLLHMTALSVERYLAICRPLRARV 148
Db 89 LGMPLEIYEMWHNYPFLFGPGVCFKALTETVCFASILSVTTVSVERYVAIVHPFRAKL 148
Qy 149 LVTRRRVRALLAIVLWVALLSAGFFLPLGV--EQDPGISVVPLNGLTARIASPLASSP 206
Db 149 ESTRRRALRYLSLVGSFVSFLPTSNIHGKFQHPNGSSVPG-----SATCTVTK 200
Qy 207 PLWSRAPPPSPGPTABAAALFSRECRPSPAQLGALRVMLVTWTAYPF-LPFLCLSI 265
Db 201 PMWV-----YNLIQNTSFPLYLPMTLISV 226
Qy 266 LYGLIGREL 274
Db 227 LYVLMLGLRL 235

RESULT 15
Q9ESQ4 PRELIMINARY; PRT; 395 AA.

ID Q9ESQ4 AC Q9ESQ4 DT 01-MAR-2001 (TREMBLrel_16, Created)
DT 01-MAR-2001 (TREMBLrel_16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel_23, Last annotation update)
DE G protein-coupled receptor TGR-1.
GN TGR-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX [1]
SEQUENCE FROM N.A.
RP MEDLINE=20449029; PubMed=10887190;
RX Heooya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani J., Fukushima S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neuromedin U Receptor";
RL J. Biol. Chem. 275:29528-29532(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AB041229; BAB13722.1; -
DR InterPro; IPR000276; GPCR Rhodpsn.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005392; NeuromedinU2.
DR Pfam; PF00001; 7cm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01567; NEUROMEDINU2.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44756 MW; EABD6D36E9C355DA CRC64;

Query Match 16.1%; Score 329; DB 11; Length 395;

Best Local Similarity 30.1%; Pred. No. 6.4e-18;					
Matches 75; Conservative 55; Mismatches 73; Indels 46; Gaps 5;					
Qy	30	CSPEPLGALYPTAVCLCLFVGVSGNVVTVMLIGRYDMRTTNLYLGSMAVSDLLI-L	88		
Db	29	CGPKRSDLSLFVSAYALIFLVGVGNLLVCMVIVRHQTLKTPNTNYLFLSLAVSDLLVLL	88		
Qy	89	LGLPFDLYRLWRSPWVFGPLLCRLSLYVGEGCTYATLLHMTALSVERYLAICRPLRARV	148		
Db	89	LGMPLIYENWHNYPFVFGVGCYFKTALFETVCFASILSVTVTSVERYVAIVHPPRAKL	148		
Qy	149	LVTRRVRVALIAVMAVALLSAGEFLFVGV--EQDPGISVWPGINGTARIASPLASSP	206		
Db	149	ESTRRRALRILSLVWSPSVFSLPNTSIHGKIFQHPNGSSVPG-----SATCTVTK	200		
Qy	207	PLMLSRAPPPSPSGPETAEEAALFSRECRPSPAQLGALRAVLMVWTYFF-LPFLCLSI	265		
Db	201	FMWV-----YNLIIQATSFIFYILPMTLISV	226		
Qy	266	LYGLIGREL	274		
Db	227	LYYLMGLRL	235		

Search completed: January 1, 2004, 06:22:27
Job time : 64.3659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 14.4561 Seconds
(without alignments)
1340.261 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDPGEGAREPPWP.....DTGGDTVGYTETSANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2155	100.0	412	1	MTLR_HUMAN
2	864.5	40.1	366	1	GHRSR_PIG
3	863.5	40.1	366	1	GHRSR_HUMAN
4	861.5	40.0	364	1	GHRSR_RAT
5	614	28.5	257	1	GHRSR_MOUSE
6	501	23.2	424	1	NTR1_RAT
7	499.5	23.2	424	1	NTR1_MOUSE
8	497	23.1	418	1	NTR1_HUMAN
9	413	19.2	410	1	NTR2_HUMAN
10	404	18.7	417	1	NTR2_MOUSE
11	401	18.6	416	1	NTR2_RAT
12	370.5	17.2	362	1	SSR5_MOUSE
13	359	16.7	363	1	SSR5_RAT
14	357.5	16.6	418	1	SSR3_HUMAN
15	355.5	16.5	453	1	GP39_HUMAN
16	344	16.0	364	1	SSR5_HUMAN
17	343.5	15.9	380	1	OPRK_CAVPO
18	342.5	15.9	370	1	OPRK_CAVPO
19	339	15.7	380	1	OPRK_RAT
20	335	15.5	380	1	OPRK_HUMAN
21	335	15.5	388	1	SSR4_HUMAN
22	334.5	15.5	367	1	OPRK_MOUSE
23	334.5	15.5	367	1	OPRK_RAT
24	334.5	15.5	370	1	OPRK_PIG
25	334.5	15.5	428	1	SSR3_RAT
26	333.5	15.5	372	1	OPRD_HUMAN
27	332.5	15.4	395	1	TRFR_CHICK
28	332	15.4	372	1	GALS_RAT
29	331.5	15.4	428	1	SSR3_MOUSE
30	330.5	15.3	370	1	OPRK_HUMAN
31	329	15.3	380	1	OPRK_MOUSE
32	327.5	15.2	371	1	GALS_MOUSE
33	327.5	15.2	384	1	SSR4_RAT
					043193 homo sapien
					Q95254 sus scrofa
					Q92847 homo sapien
					O08725 rattus norv
					O99p50 mus musculus
					P20789 rattus norv
					O68319 mus musculus
					P30989 homo sapien
					O95665 homo sapien
					P70310 mus musculus
					O63394 rattus norv
					O08858 mus musculus
					P30938 rattus norv
					P32745 homo sapien
					O43194 homo sapien
					P35346 homo sapien
					P41144 cavia porce
					P47748 rattus norv
					P34975 cavia porce
					P41145 homo sapien
					P31391 homo sapien
					P35377 mus musculus
					P35370 rattus norv
					P79292 sus scrofa
					P30936 rattus norv
					P41143 homo sapien
					O93603 gallus gall
					O08726 rattus norv
					P30935 mus musculus
					P41146 homo sapien
					P33534 mus musculus
					O88854 mus musculus
					P30937 rattus norv

RESULT 1
MTLR_HUMAN STANDARD; PRT; 412 AA.
AC O43193;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Motilin receptor (G protein-coupled receptor GPR38).
GN GPR38 OR MTLR1 OR MTLR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=98110578; PubMed=9441746;
RA McKee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
RT "Cloning and characterization of two human G protein-coupled receptor
genes (GPR38 and GPR39) related to the growth hormone secretagogue
RT and neuropeptide receptors";
RL Genomics 46:426-434(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=99316084; PubMed=10381885;
RA Feighner S.D., Tan C.P., McKee K.K., Palyha O.C., Hreniuk D.L.,
RA Pong S.-S., Austin C.P., Figueroa D., MacNeil D., Cascieri M.A.,
RA Nargund R., Bakshi R., Abramovitz M., Stocco R., Kargman S.,
RA O'Neill G., van der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,
RA Howard A.D.;
RT "Receptor for motilin identified in the human gastrointestinal
RT system";
RL Science 284:2184-2188(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RA Wall M.;
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION.
RX MEDLINE=21219832; PubMed=11322507;
RA Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,
RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
RT "Growth hormone secretagogue receptor family members and ligands";
RL Endocrine 14:9-14(2001).
CC -1- FUNCTION: Receptor for motilin.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=O43193-1; Sequence=Displayed;
CC Name=B;
CC IsoId=O43193-2; Sequence=VSP_001894;
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE
CC MARROW.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----

ALIGNMENTS

34	326	15.1	519	1	TLR2_DROME	P30375 drosophila
35	325	15.1	453	1	CCXR_XENLA	P70031 xenopus lae
36	322.5	15.0	370	1	GALT_RAT	O88626 rattus norv
37	322	14.9	384	1	SSR4_MOUSE	P49660 mus musculus
38	321.5	14.9	372	1	OPRD_RAT	P33533 rattus norv
39	321.5	14.9	398	1	TRFR_HUMAN	P14981 homo sapien
40	321	14.9	353	1	APJ_XENLA	P79960 xenopus lae
41	320	14.8	372	1	OPRD_MOUSE	P32300 mus musculus
42	319.5	14.8	387	1	GALS_HUMAN	O43603 homo sapien
43	319.5	14.8	398	1	TRFR_BOVIN	O46639 bos taurus
44	318.5	14.8	368	1	GALT_HUMAN	O60755 homo sapien
45	318.5	14.8	398	1	TRFR_SHEEP	Q28596 ovis aries

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF034632; AAC26081.1; -;
EMBL; AL137000; CAC19107.1; -;
Genew; HGNC:4495; GPR38.
DR MTM; 602885; -;
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0007586; P: digestion; TAS.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . . ; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 56 1 (POTENTIAL).
FT DOMAIN 57 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 94 2 (POTENTIAL).
FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 134 3 (POTENTIAL).
FT DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 270 5 (POTENTIAL).
FT DOMAIN 271 298 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 299 320 6 (POTENTIAL).
FT DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 335 358 7 (POTENTIAL).
FT DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 235 BY SIMILARITY.
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 301 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
LVVLAFLIICWLPFHVGRIRIYINTEDSRMMYFQYFNIVALQFLYSASINPILYNLISK
QLFYLASINPILYNLISKYRAAFKLLARKSRPRGFRH
SRDTAGVAGDGTGVGTETGSAVNKTMG -> RKWSRRG
SKDAQLQSPAGTAQTLGPIPLLAQLWAPLPAPPTIPAS
TRGGSGIYNLVALPRWQVHLKGRFADVLLSVL
(in isoform B).
/FTID=VSP_001894.

SQ SEQUENCE 412 AA; 45344 MW; C13FP6165012DEF3 CRC64;
Query Match 100.0%; Score 2155; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 6e-137;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGEGAREPPALPCDERRCSPPLGALVPTAVCLFLVVGSGNVTV 60
DB 1 MGSPWNGSDGEGAREPPALPCDERRCSPPLGALVPTAVCLFLVVGSGNVTV 60
QY 61 MLIGRYDMRTTNLYLGSAVSDLLILLGLPDLVLMRSRPWFGPLLCRLSLVYVGE 120
DB 61 MLIGRYDMRTTNLYLGSAVSDLLILLGLPDLVLMRSRPWFGPLLCRLSLVYVGE 120
QY 121 CTYATLLHMTALSVERVLAICRPLARVLVTRVRVLAIVLWAVALLSAGPFLVGV 180
DB 121 CTYATLLHMTALSVERVLAICRPLARVLVTRVRVLAIVLWAVALLSAGPFLVGV 180
QY 181 QDFGISVVGINGTARIASSPLWLSRAPPPSPGPTAAALFSECRPSPA 240
DB 181 QDFGISVVGINGTARIASSPLWLSRAPPPSPGPTAAALFSECRPSPA 240
QY 241 QLGAALRVMLVTTAYFPFLPCLSLYGLIGRLWSSRRPLRGPASGRGRHQTVRVL 300
DB 241 QLGAALRVMLVTTAYFPFLPCLSLYGLIGRLWSSRRPLRGPASGRGRHQTVRVL 300

QY 301 LVVLAFLIICWLPFHVGRIRIYINTEDSRMMYFQYFNIVALQFLYSASINPILYNLISK 360
DB 301 LVVLAFLIICWLPFHVGRIRIYINTEDSRMMYFQYFNIVALQFLYSASINPILYNLISK 360
QY 361 KYRAAFAKLLARKSRPRGFRHSDRTAGVAGDGTGVGTETGSAVNKTMG 412
DB 361 KYRAAFAKLLARKSRPRGFRHSDRTAGVAGDGTGVGTETGSAVNKTMG 412

RESULT 2

GHSR_PIG
ID GHSR_PIG STANDARD; PRT; 366 AA.
AC Q95254; Q95255;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).
GN GHSR.
OS Sus acrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
OX [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RP STRAIN=Yorkshire; TISSUE=Pituitary;
RX MEDLINE=96337998; PubMed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P., Liberato P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L., Palyha O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K., McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevich M., Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G., Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K., Schaeffer J.W., Smith R.G., van der Ploeg L.H.T.;
RA "A receptor in pituitary and hypothalamus that functions in growth hormone release.";
RT Science 273:974-977(1996).
RL
CC -I- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1A;
CC IsoId=Q95254-1; Sequence=Displayed;
CC Name=1B;
CC IsoId=Q95254-2; Sequence=VSP_001918, VSP_001919;
CC -I- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U60178; AAC48630.1; -;
EMBL; U60180; AAC48631.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 96 2 (POTENTIAL).

```

FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVVFAPILCWLPFHVGRYLFSSK -> GGSQCALELSLPG.
FT VARSPLIC 266 289 PLHSSCLFSSP (in isoform 1B).
FT VARSPLIC 290 366 /FTID=VSP 001918.
FT VARSPLIC 290 366 Missing (in isoform 1B).
FT SEQUENCE 366 AA; 41194 MW; 2C850B3BF61B7C1C CRC64;
Query Match 40.1%; Score 864.5; DB 1; Length 366;
Best Local Similarity 45.1%; Pred. No. 5.9e-51;
Matches 189; Conservative 55; Mismatches 104; Indels 71; Gaps 10;

QY 5 WNGSDGEGA-----REPPWALPPCD---ERRCSPPLGALVPVTVAVCLCLFVWVSGN 56
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 WNATPSEPGPNLTLPDGLWDAPPENDSLVEELLPLFPTLLAGVATATCVALFVWVGIAGN 61
QY 57 VVTVMILGRVDRMTTNLYLGSMAVSDILILGLPDLRLWRSRPWVGPPLICRLSLY 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 LLTMLVRSRPREMTTNLYLGSMAVSDILILGLPDLRLWRSRPWVGPPLICRLSLY 121
QY 117 VGGCTVATLLHTALSVRYLAICRLRLRVLRVTRRRVALLAVMAVALLSAGPPLFL 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 VSSECTVATVLTALSVRYFAICPLRAKVVTGKRVKLVILVIAVAFCSAGPFLVL 181
QY 177 VGVEQDVGISVWPLNGLNTARIASPLASPPPLMLSRAPPSPGPTABAAALFSRECR 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
182 VGVEHD-----NCT-----DPRD-----TWECR 199
QY 237 PS--PAQLGALRVMLWTTAYFFLPFLCLILYGLIGRELWSSRRPLRGPAAAG---RER 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 ATEFAVRSGLTVMVWVSSVFFFLVFLVCLVLYSLIGRLKLM---RRKRGEAAVGSLLRDQ 256
QY 292 GHRQTVRLVAVVLAFLICLPLPHVGRIV---INTEDSRMVFSQYFNIVALQLFVLSA 348
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 NHQTVRLVAVVLAFLICLPLPHVGRIVFSKSLPESGVEIAQISQYCNLVSVFLFVLSA 316
QY 349 SINPILYNLSKKYRAAFAKLLARKSRPRGFRHSRDTAGEVAGDTGGDTGVGTETSAN 407
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 AINPILYNLSKKYRAVAVFKLLGPEPFSQKLSLTKDESSR-----ANTESSIN 365

RESULT 3
GHSR HUMAN
ID GHSR_HUMAN STANDARD; PRT; 366 AA.
AC Q92847; Q92848; Q96R37;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (Ghrelin receptor).
GN GHSR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RC TISSUE=Plutitary;
RX MEDLINE=96337998; PubMed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,

```

```

RA Liberato P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
RA Palyha O.C., Anderson J., Pares S.S., Diaz C., Chou M., Liu K.K.,
RA McKee K.K., Pong S.-S., Chung L.-Y., Elbrecht A., Dashkevich M.,
RA Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,
RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaffer J.M., Smith R.G., van der Ploeg L.H.T.,
RT "A receptor in pituitary and hypothalamus that functions in growth
RT hormone release.";
RL Science 273:974-977(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RX MEDLINE=21255649; PubMed=11356716;
RA Petersen S., Rasch A.C., Penschorn M., Bell F.U., Schulte H.M.;
RT "Genomic structure and transcriptional regulation of the human growth
RT hormone secretagogue receptor.";
RL Endocrinology 142:2649-2659(2001).
RN [3]
RP FUNCTION.
RX MEDLINE=21219832; PubMed=11322507;
RA Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,
RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
RT "Growth hormone secretagogue receptor family members and ligands.";
RL Endocrine 14:9-14(2001).
RN [4]
RP FUNCTION.
RX MEDLINE=20087959; PubMed=10604470;
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
RT "Ghrelin is a growth-hormone-releasing acylated peptide from
RT stomach.";
RL Nature 402:656-660(1999).
CC -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1A;
CC IsoId=Q92847-1; Sequence=Displayed;
CC Name=1B;
CC IsoId=Q92847-2; Sequence=VSP_001916, VSP_001917;
CC -!- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).

```

```

FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVVFATLWLPVHGVYFSKS -> GGSQALRLSLAG
PILSLCLPSL (in isoform 1B).
/FTID=VSP_001916.
FT VARSPLIC 290 366 Missing (in isoform 1B).
/FTID=VSP_001917.
SQ SEQUENCE 366 AA; 41328 MW; D1B62710DA9DC0C6 CRC64;

Query Match 40.18; Score 863.5; DB 1; Length 366;
Best Local Similarity 44.78; Pred. No. 6.9e-51;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGSGN 56
DB 2 WNATPSEPGFNLTDLDDWDASPGNDSLGDELLQLFPAPLLAGVATCTVALFVVGAGN 61

QY 57 VVTVMILGRYDRMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSLY 116
DB 62 LLTGLVTVSRPRELRTTNLYLSSMAFSDLLIFLCMLDPLVRLQYRPWNFGDLCKLFQF 121

QY 117 VGGCTATLHWTALSVRYLAICRPLRVLVTRRRVALLAVLWALLSAGFLFL 176
DB 122 VBSCTATVLTITALSVERFYAFCFLRAKVVTGKRVLFVFWAVAFCSAGIFVL 181

QY 177 VGYEODPGISVWPGCLNTARIASSPLASPPMLSRAPPSPPGPTAEAAALFSGREC 236
DB 182 VGYEHE-----NGT-----DP-W-----DTNECR 199

QY 237 PS--PAQLGALRVLMTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAASGRGRHR 294
DB 200 PTEFAVRSGLTVMWVSSIFFFLPVFLVFLVLSLIGRLWRRRGDAVVGASLRDQNHK 259

QY 295 QTVRVLLVVLVLAFLIICWLPVHGVRIIYNTEDS---RMVFSQYFNIVALQLPYLSASIN 351
DB 260 QTVKMLAVVVFATILCWLPHVGVYFSKSFEPGSLKIAQISQYCNLVSFVFLYLSAAN 319

QY 352 PILYNLISKYRAAFAKLLARKSRPRGPHRSRDTAGEVAGDTGDTVGYTETSAN 407
DB 320 PILYNLISKYRVAVFLLGFEFPFSQKLSLTKDESSR-----AWTESSIN 365

RESULT 4
GHSR_RAT STANDARD; PRT; 364 AA.
AC Q08725;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (Ghrelin receptor).
GN GHSR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=97246555; PubMed=9092793;
RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C.P.,
Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;
"Molecular analysis of rat pituitary and hypothalamic growth hormone
secretagogue receptors.";
Mol. Endocrinol. 11:415-423 (1997).
[2]
RN SEQUENCE OF 1-240 FROM N.A.
RC STRAIN=Wistar; TISSUE=Pituitary;
RX MEDLINE=98100386; PubMed=9437732;
RA Yokote R., Sato M., Matsubara S., Obye H., Niimi M., Murao K.,
Takahara J.;
"Molecular cloning and gene expression of growth hormone-releasing
peptide receptor in rat tissues.";
Peptides 19:15-20 (1998).
[3]
RN FUNCTION.
RX MEDLINE=20067959; PubMed=10604470;
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
"Ghrelin is a growth-hormone-releasing acylated peptide from
stomach.";
Nature 402:656-660 (1999).
RL -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine) (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U94321; AAC53156.1; -
DR EMBL; AB001982; BAA21777.1; ALT INIT.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 73 96 2 (POTENTIAL).
FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 303 326 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 327 364 BY SIMILARITY.
FT DISULFID 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 40963 MW; DCF559BE061EE9 CRC64;

Query Match 40.0%; Score 861.5; DB 1; Length 364;
Best Local Similarity 48.3%; Pred. No. 9.3e-51;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGPEGAREP-----PWPALPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGSG 55
DB 2 WNATPSEPGFNVTLDLDDWDASPGNDSLGDELLQLFPAPLLAGVATCTVALFVVGISG 59

QY 56 NVVTVMILGRYDRMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSL 115
```



```

Db 60 NLLTMLVSRPRELTTNLYLSMAFSDLLIFCLMPLDLVRLWQYRPMWFGDLLCKLFQ 119
Qy 116 YVGGCTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLALVAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYFAICPLRAKVVTKGRVKLVILVAVAFSAGPIFV 179
Qy 176 LVGVEQDGISSVVGELNGTARIASSPLASSPPLWLSRAPPPSPGTPETAAALPFRRC 235
Db 180 LVGVEH- - - - -NGT-----DPRD-----TNEC 197
Qy 236 RPS--PAQLGALRVMLVMTTAYFPFLCLSLYLGLIGRELWSSRRPLRGPASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVCLTVLSLIGRLKLV--RR--RGDAVAGSLRD 253
Qy 291 RGHRTVRLVLLVLAFLICWLPVHGVRIIYNTEDS---RMVFSQYFNIVLQLFVLS 347
Db 254 QNHKQTVKMLAVVFAFLCWLPHVGRVLFKSFEPGSLIAQISQYCNLVSPVLFVLS 313
Qy 348 ASINPILYNISKYRAAFKLL 370
Db 314 AAINPILYNIMSKYRVAVFKLL 336

```

RESULT 5

```

ID GHRS MOUSE STANDARD; PRT; 257 AA.
AC Q99P50; Q91282;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (Ghrelin receptor) (Fragment).
GN GHRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE OF 1-183 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RA Kacsob B.;
RT "Cloning of mouse ghrelin/growth hormone secretagogue receptor cDNA by
RT rapid amplification of cDNA ends (RACE).";
RL Submitted (GSP-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE OF 73-257 FROM N.A.
RC STRAIN=129S3/SvimJ;
RA Peng X., Frohman L.A., Kineman R.D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine) (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AY056474; AAL13336.1; -
CC EMBL; AF323997; AAG61141.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.

```

```

KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 35
FT TRANSMEM 36 58
FT DOMAIN 59 77
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 100
FT DOMAIN 101 125
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 148
FT DOMAIN 149 160
FT TRANSMEM 161 183
FT DOMAIN 184 211
FT TRANSMEM 212 234
FT TRANSMEM 235 257
FT DISULFID 115 197
FT CARBOHYD 13 13
FT CARBOHYD 26 26
FT CARBOHYD 187 187
FT NON TER 257 257
SQ SEQUENCE 257 AA; 28743 MW; D6PA21EAB0E30791 CRC64;

```

Query Match 28.5%; Score 614; DB 1; Length 257;
 Best Local Similarity 45.1%; Pred. No. 2.1e-34;
 Matches 137; Conservative 39; Mismatches 66; Indels 62; Gaps 10;

```

Qy 5 WNGSDGPEGAREP-----PWPALPPCD-----ERRCSPPFLGALVPVTAVCLELVWCVSG 55
Db 2 WNAT--PSEPEPNVTLDLDWDASPGNDSLSDELPLFPAPLLAGVTATCAVLFVWGISS 59
Qy 56 NVVTMLIGRYDMRTTNLYLGSMVAVSDLLILGLPFDLYRLWRSRPMWFGDLLCKLSL 115
Db 60 NLUTMLVSRPRELTTNLYLSMAFSDLLIFCLMPLDLVRLWQYRPMWFGDLLCKLFQ 119
Qy 116 YVGGCTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLALVAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYFAICPLRAKVVTKGRVKLVILVAVAFSAGPIFV 179
Qy 176 LVGVEQDGISSVVGELNGTARIASSPLASSPPLWLSRAPPPSPGTPETAAALPFRRC 235
Db 180 LVGVEH- - - - -NGT-----DPRD-----TNEC 197
Qy 236 RPS--PAQLGALRVMLVMTTAYFPFLCLSLYLGLIGRELWSSRRPLRGPASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVCLTVLSLIGRLKLV--RR--RGDAVAGSLRD 253
Qy 291 RGHRTVRLVLLVLAFLICWLPVHGVRIIYNTEDS---RMVFSQYFNIVLQLFVLS 347
Db 254 QNHKQTVKMLAVVFAFLCWLPHVGRVLFKSFEPGSLIAQISQYCNLVSPVLFVLS 313

```

RESULT 6

```

ID NTR1 RAT STANDARD; PRT; 424 AA.
AC P20789;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity levocabastine-
DE insensitive neurotensin receptor) (NTRH).
GN NTR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90297956; PubMed=1694443;
RA Tanaka K., Masu M., Nakanishi S.;
RT "Structure and functional expression of the cloned rat neurotensin
RT receptor.";
RL Neuron 4:847-854(1990).
CC -1- FUNCTION: RECEPTOR FOR THE TRIPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.

```

CC	- -	SUBCELLULAR LOCATION: Integral membrane protein.
CC	- -	SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC		HIGHEST TO TACHYKININS RECEPTORS.
PIR,	JH0164;	JH0164.
DR	InterPro:	IPR000276; GPCR_Rhodpsn.
DR	InterPro:	IPR003985; NTL_rec.
DR	InterPro:	IPR003984; NT_rec.
DR	Fram:	PF00001; 7tm_1; 1.
DR	PRINTS:	PR00237; GPCRRHOOPSN.
DR	PRINTS:	PR01479; NEUROTENSINR.
DR	PRINTS:	PR01480; NEUROTENSINR.
DR	PROSITE:	PS00237; G PROTEIN RECP_F1_1; 1.
DR	PROSITE:	PS00262; G_PROTEIN_REC_F1_2; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;	
KW	Phosphorylation; lipoprotein; Palmitate.	
FT	DOMAIN 1	64 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	65 1 (POTENTIAL).
FT	DOMAIN	88 96 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	97 121 2 (POTENTIAL).
FT	DOMAIN	122 143 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	144 165 3 (POTENTIAL).
FT	DOMAIN	166 188 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	189 210 4 (POTENTIAL).
FT	DOMAIN	211 235 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	236 260 5 (POTENTIAL).
FT	DOMAIN	261 308 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	309 330 6 (POTENTIAL).
FT	DOMAIN	331 348 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	349 372 7 (POTENTIAL).
FT	DOMAIN	373 424 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	142 225 BY SIMILARITY.
FT	LIPID	388 388 PALMITATE (POTENTIAL).
SEQ	SEQUENCE	424 AA; 47054 MW; A9C2F7EAF8D9BCD3 CRC64;
Query Match		
Best Local Similarity 32.2%; Score 501; DB 1; Length 424;		
Matches 125; Conservative 59; Mismatches 110; Indels 86; Gaps 11;		
QY	39	VPTAVCLCLFFVGVGSGNVTVMLGR---YRDRTTNLYLGSMVSDLLI-LLGLPFD 94
DB	65	VLVTAIYLALFVVGTGVGNSVTAFTLARKKLSQLSTVHVHLGSLASDLLILLAMPVE 124
QY	95	LXR-LARSRPWPFGPLL-CRLSLYVGEGCTATLLHMTALSVERVLAICRPLRARVLTNR 153
DB	125	LYNFIVWHHPWAGDGRCGYFLRDCACTATALNVASLSVERVLAICHFPKAKTLMRSR 184
QY	154	RVRALLIAVLMAVALLSAGPPLFLVGVBDQPGISVWPGINGTARIASSPLASSPPWLSSRA 213
DB	185	RTKKFISAIWLASALLAIPMLFTMGLQNRSRGDTHFG-----GLVCTPIVDT----- 231
QY	214	PPSPSPGPBETAABAAALFSRECRPSAQALGARVMLWTTAYPFL-PFLCLSLYGLIGR 272
DB	232	-----ATVKVIQVNTFMSELPFMLVISINTLVIAN 262
QY	273	ELWSSRRPLRGPAASGR-----ERGHQTNR-----VLLVVVFALFI 309
DB	263	KLTVM---VHQAAEQQRVCVTGTHNGLEHSTFNNTIEPGRVQALRHGVLVLRVAVVIAFV 319
QY	310	CMLPFPHVGRIIYINTEDSR---MMVFSQYFNIVALQFLYSASINFILYNLISKKYRAA 365
DB	320	CMLPYHVRMLMFCYISDEQWTTFDFDYHHVYMLTNALFVSSAINFILNLVSANFRQV 379
QY	366	AFKLLIA-----RKSRP 377
DB	380	FLSTLACLCPGWRHRKKRP 399
RESULT 7		
ID	NTRI_MOUSE	STANDARD; PRT; 424 AA.
ID	NTRI_MOUSE	

088319;
16-OCT-2001 (Rel. 40, Created)
DT
16-OCT-2001 (Rel. 40, Last sequence update)
DT
16-OCT-2001 (Rel. 40, Last annotation update)
DE
Neurotensin receptor type 1 (NT-R-1).
DE
NTR1 OR NTSR
GN
Mus musculus (Mouse).
OS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC
NCBI_TaxID=10090;
OX
[1]
RN
SEQUENCE FROM N.A.
RP
TISSUE=Brain;
RC
Snider J., Sano H., Ohta M.;
RT "Neurotensin receptor type 1";
RA Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RL -!- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; AB017027; BAA33013.1; -;
DR MGD; MG1:97386; Ntsr.
DR InterPro; IPR000276; GPCR Rhodspn.
DR InterPro; IPR003985; NT1_rec.
DR InterPro; IPR003984; NT_rec.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODSPN.
DR PRINTS; PR01479; NEUROTENSINR.
DR PRINTS; PR01480; NEUROTENSIR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS02662; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein;
KW phosphorylation; Lipoprotein; Palmitate.
DR DOMAIN 1 64
FT EXTRACELLULAR (POTENTIAL).
FT
FT TRANSMEM 65 87
FT DOMAIN 88 96
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 97 121
FT DOMAIN 122 143
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 144 165
FT DOMAIN 166 188
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 189 210
FT DOMAIN 211 235
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 236 260
FT DOMAIN 261 308
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 309 330
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 331 348
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 349 372
FT DOMAIN 373 424
FT CARBOHYD 4 4
FT FT CARBOHYD 38 38
FT FT CARBOHYD 42 42
FT FT CARBOHYD 211 211
FT DISULFID 141 224
FT LIPID 388 388
SQ SEQUENCE 424 AA; 47216 MW; 8E9A723171A48711 CRC64;

Query Match 23.2%; Score 499.5; DB 1; Length 424;
Best Local Similarity 33.3%; Pred.No. 1.5e-26;
Matches 126; Conservative 57; Mismatches 114; Indels 81; Gaps 11;

39 VPVTAVCLCLFVGVSQGVNVTVMLIGR---YDMRETTTNLYLGSMVSDLLI-LGLGPF 94

64	VLVTAVYALFVUGTVGNSVTAFTLARKKSQSQSTQVHYHGLSIALSDLLILLAMPVE	123
Qy	LYR-LRMSRPVFGFLLCRLSLYVGEGETYATLLHMTALSVERYIAICRPLRARVLVTRR	153
Db	LYNFIWVHPWAFGDACRGYFELRDACYTATNALNVASISVERYIAICHPFKAKTLMGRS	183
Qy	RVRALIIVNAVALLSAGPFPLFVGVBEODPGISVVVPGLINGTARIASSPLASSPPLWLSRA	213
Db	RTKKFISAIWLASALLAVPMLFTWGLQ-----NRS	213
Qy	PPSPSPSGPETAEEAALFSCRSPSPAQGLALRWMLWTTTAYPFL-PPFLCSILYGLIGR	272
Db	ADGQHPG-----LVCTPT-VDTATVKVVIQNTFMSFLFPMLIISILNTVIAN	261
Qy	EL-----WSSRRRLRGPAA-----SGRERGHRTQVRVLLVVVLAFLICW	311
Db	KLTVMVHQAEQGRGVCTVGTGTHNSLEHSTFNMSIEPGRVQALRHGVLVRVAVIIVCW	321
Qy	LPFHVGRIIYINTDSR-----MMYFSQYFNIVALQFYLSSINPILYNLISKKYRAAAF	367
Db	LPYHVRRLMFCYISDEQWTTLEDFYHYFMYLMTNALFYVSSAINPILYNLVSANFQVFL	381
Qy	KLLLA-----RKSRP	377
Db	STLACLPGWRRRRRKRP	399

RESULT 8

NTR1_HUMAN	NTR1_HUMAN	STANDARD;	PRT;	418 AA.
ID	P09889; Q9H4H1; Q9H4T5;			
AC	F30989; Q9H4H1; Q9H4T5;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neurotensin receptor type 1 (NT-R-1) (High-affinity levocabastine-insensitive neurotensin receptor) (NTRH).			
GN	NTR1 OR NTRR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93154505; PubMed=8381365;			
RX	Vita N., Laurent P., Lefort S., Chalou P., Dumont X., Kaghad M.,			
RX	Gully D., le Fur G., Fetrara P., Caput D.;			
RT	"Cloning and expression of a complementary DNA encoding a high			
RT	affinity human neurotensin receptor.";			
RL	FEBS Lett. 317:139-142(1993).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RX	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.R., Heath P.D., Ho S.D., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.W., Johnson D.			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.			
RA	Lehvaestaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,			
RA	Skuce C.D., Smith M.B., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Syncamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			

Db 64 VLTVAVLALFVVGVTVGNVTVAFTLARKKSLQSLQSTVHYHLSLALSDLLTLLAMPVE 123
 QY 95 LYR-LWRSRPWVFGPCLLCRLSLVYGEVGTATLHMTALSVERYLACRPLARVLVTRR 153
 Db 124 LYNFVWHHPWAFGDACRGYFYRDACATATANVASLVERYLACHFFKATLMSRS 193
 QY 154 RVRLIALVLAVALLSAGPFLFLVGVQDPGISVVPGLNGTARTIASSPLSLMSRA 213
 Db 184 RTKKFISAIWLASALLVPMFLTMG-EQN----- 211
 QY 214 PPSPPSGPTAEAAALFSECRSPQAQGLRVLMLWVTT-AVFFLPCLCLSLYGLIGR 272
 Db 212 -----RSADQHQAGGLVCTPT-IHTATVKVVIQVNTFMSPFPMVSVLNTIIAN 261
 QY 273 ELWSSRRPL--RGAAS-----GREGRHQRTVRLVLLVVLAFIICWLPFHV 316
 Db 262 KLTVMVQAQEGQVCTVGEHSTFSMAIEPGRVQALRHGVRVLRVAVVAVVWCWLPFHV 321
 QY 317 GRIIYINTDSR-----MMY-FSQYFNIVALQIFYLASINFLYNLISKYR-----AA 365
 Db 322 RRLMFCVISDEQWTFELYDFHYFMYVTNALFYVYSTINFLYNLVSANFRHIFLATLAC 381
 QY 366 AFKLLARKSRPGRFHSRDT 386
 Db 382 LCPVWRRRRKRP-AFSRKADS 401

RESULT 9

NR2 HUMAN STANDARD; PRT; 410 AA.
 AC 095665; Q8TBH6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neurotensin receptor type 2 (NT-R-2) (Levocabastine-sensitive
 DE neurotensin receptor) (NTR2 receptor).
 GN NTR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99066919; PubMed=9851594;
 RA Vita N., Oury-Donat F., Chalou P., Guillemot M., Kaghad M., Bachy A.,
 Thurnissen S., Garcia S., Poinot-Chazel C., Casellas P., Keane P.,
 Le Fur G., Maffrand J.-P., Shoubrie P., Caput D., Ferrara P.;
 RT "Neurotensin is an antagonist of the human neurotensin NT2 receptor
 expressed in Chinese hamster ovary cells.";
 RL Eur. J. Pharmacol. 360:265-272(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
 Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heish F.,
 Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smallus D.E.,
 Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -I- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHIKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; Y10148; CAAY1233.1; -.
 CC EMBL; BC022501; AAH22501.1; -.
 CC Genew; HGNC:8040; NTR2.
 CC MIM; 605538; -.
 CC GO; GO:0005887; C: integral to plasma membrane; TAS.
 CC GO; GO:0007600; P: sensory perception; TAS.
 CC InterPro; IPR00276; GPCR Rhodopsin.
 CC InterPro; IPR003986; NT2 rec.
 CC InterPro; IPR003984; NT rec.
 CC Pfam; PF00001; 7tm_1_1.
 CC PRINTS; PR00237; GPCR RHODOPSIN.
 CC PRINTS; PR01479; NEUROTENSINR.
 CC PRINTS; PR01481; NEUROTENSINR.
 CC PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
 CC PROSITE; PS00262; G PROTEIN RECEPTOR FL 2; 1.
 CC G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 CC DOMAIN 1 32
 CC TRANSMEM 33 55
 CC DOMAIN 56 64
 CC TRANSMEM 65 87
 CC DOMAIN 88 109
 CC TRANSMEM 110 131
 CC DOMAIN 132 154
 CC TRANSMEM 155 176
 CC DOMAIN 177 217
 CC TRANSMEM 218 237
 CC DOMAIN 238 297
 CC TRANSMEM 298 318
 CC DOMAIN 319 337
 CC TRANSMEM 338 358
 CC DOMAIN 359 410
 CC DISULFID 108 194
 CC LIPID 377 377
 CC CONFLICT 168 168 L -> M (IN REF. 2).
 CC CONFLICT 367 367 K -> R (IN REF. 2).
 CC SEQUENCE 410 AA; 45413 MW; 8C3ADA22BE15FD66 CRC64;
 Query Match 19.2%; Score 413; DB 1; Length 410;
 Best Local Similarity 31.4%; Pred. No. 8.3e-21;
 Matches 130; Conservative 60; Mismatches 134; Indels 90; Gaps 16;
 QY 17 PMPALPP---CDERCSPPFPFALVPTAVCLFLFVGVSGNVVTMLI-----GR 65
 Db 8 PPRPSNPGLSLDARLGVDTRLWAKVLTALYALWALGAAGNALSVHVVVKARAGR 67
 QY 66 YRDMRTTNYLGSMASVLDLLILL-GLPPDIYR-LWRSRPWVFGPCLLCRLSLVVGECY 123
 Db 68 LRH-----HVLSTALAGLLLLVGVFVVELYSFVFWFVDFGLGCRGYVHELCA 120
 QY 124 ATLLHMTALSVRYLAICRPLARVLVTRRRVRLIALVLAVALLSAGPFLVGV---- 179
 Db 121 ATVLVAGLSAERCLAVCOPLRARSLLTPRTFWLVALSWAASLGIALPNAVINGQKHEL 180
 QY 180 -----EODPGISVVPGLNGTARTIASS-----PLASSPPI-----WLS 211
 Db 181 ETADGGEPEASRVCITVL--VSR TALQVFIQVNVVLSFVLPLALTAFNGVTVSHLLALCS 238

CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER
CC LEVELS SEEN IN THE HEART AND INTESTINE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
CC EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
CC BRAIN).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97121; CAA65787.1; -;
CC PIR; S68822; S68822.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC InterPro; IPR003986; NT2_Rec.
CC InterPro; IPR003984; NT_Rec.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PRINTS; PR01479; NEUROTHENSINR.
CC PRINTS; PR01481; NEUROTHENSIN2.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
CC PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
CC G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
CC DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 33 55 1 (POTENTIAL).
CC DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 65 87 2 (POTENTIAL).
CC DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 110 131 3 (POTENTIAL).
CC DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 155 176 4 (POTENTIAL).
CC DOMAIN 177 216 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 217 237 5 (POTENTIAL).
CC DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 298 318 6 (POTENTIAL).
CC DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 338 358 7 (POTENTIAL).
CC DOMAIN 359 416 CYTOPLASMIC (POTENTIAL).
CC DISULFID 108 194 BY SIMILARITY.
CC LIPID 377 377 PALMITATE (POTENTIAL).
CC SEQUENCE 416 AA; 46265 MW; 127FC5F5CB6FE208 CRC64;
Query Match 18.68; Score 401; DB 1; Length 416;
Best Local Similarity 29.48; Pred. No. 5.3e-20;
Matches 122; Conservative 68; Mismatches 145; Indels 80; Gaps 13;
Qy 18 PWPALP-----PCDERRCSPFPGALVPVAVCLCLFVVGSGNVVTVMLIGRYDMRT 71
Db 6 PWPSPSPSAGLSLEARLGVDTRLWAKVLTALYSILFAFGTAGNLSVHVLUKARGRP 65
Qy 72 -TTNLYLGNAVSDLLILL-GLPFDLYR-LWRGRPVVFGPLCLRLSLYVGCGTYATLLH 128
Db 66 GLRLYHVLALSALLLLVSMPEMLNFVNSHYVFWFGDLGCRGGYFVRELCAATVLS 125
Qy 129 MTALSRYVLAICRLPRLVLRVRRVRLAIVLAVALLSAGFPFLPLVGV----- 179
Db 126 VASLSARCLVACQPLRRLTLRRLRLSLVWASGLALPMVIMQKHEVESADG 185
Qy 180 EODFGISVVPGLNGTARIA-----SSPLASPPPLWLSRAPPPSPGPETAEEA 228
Db 186 EPEPASVCTVLSRATLQVFIQVNLVSPALPLATFL-----NGITVNHLM 234
Qy 229 ALPSRECRPSPAQLGALRVLMTVTFPLFLCLSLYGLIGRELWSSRRPLGPAAG 288
Db 235 ALYS-QVPSAQSVPISPSLELSE-----EGLLGFITWRKTLSLGVQASLV 281

Qy 289 RE-----RHRQTVRLVLLVVLAFIICWLPFPHVGRIRIYINTED-----SRMMYFSQYFNI 338
Db 282 RHKDAQSIRLSQSAQVLRVAIVAVVYICWLPYHARLMVCIYDDGWTNELYDFYHYFM 341
Qy 339 VALQLPYLGASINPIYLNISKYRAAFAKLLARKSRPRGFRHSRDTAGEVAGD 393
Db 342 VTNTLFVSSAVTPILYNVSSFR-----KLFL-----ESLSLQGE 379

RESULT 12

SSRS_MOUSE
ID SSRS_MOUSE STANDARD; PRT; 362 AA.
AC O08858; O08998;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Somatostatin receptor type 5 (SSSR).
GN SSTR5 OR SMSTR5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Liver;
RX MEDLINE=97444289; PubMed=9300821;
RA Lublin A.L., Diehl N.L., Hochgeschwender U.;
RT "Isolation and characterization of the gene encoding the type 5 mouse
RT (Mus musculus) somatostatin receptor (msst5).";
RL Gene 195:63-66 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA Moldovan S., Demayo F., Brunicaardi P.C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Gordon D.F., Woodmansee W.W., Wood W.M., Knauf H., James R.A.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA Baumeister H., Roosterman D., Schafer J., Kreuzer O., Meyerhof W.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U82697; AAC5353.1; -;
DR EMBL; AF004740; AAB61418.1; -;
DR EMBL; AF030441; AAB86492.1; -;
DR EMBL; AF035777; AAB88302.1; ALT_INIT.
DR HSP; P02699; 1F88.
DR MGD; MGI:894282; Sstr5.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT

```
FT TRANSSEM 36 63 1 (POTENTIAL).
FT DOMAIN 64 73 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 74 99 2 (POTENTIAL).
FT DOMAIN 100 110 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 111 132 3 (POTENTIAL).
FT DOMAIN 133 154 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 155 175 4 (POTENTIAL).
FT DOMAIN 176 195 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 196 220 5 (POTENTIAL).
FT DOMAIN 221 246 6 (POTENTIAL).
FT TRANSSEM 247 272 7 (POTENTIAL).
FT DOMAIN 273 282 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 283 307 7 (POTENTIAL).
FT DOMAIN 308 362 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 109 184 BY SIMILARITY.
FT CONFLICT 99 99 V -> VV (IN REF. 1).
FT CONFLICT 303 305 YGF -> LWL (IN REF. 2).
SQ SEQUENCE 362 AA; 39948 MW; AA091DD570DFAB CRC64;

Query Match 17.2%; Score 370.5; DB 1; Length 362;
Best Local Similarity 27.8%; Pred. No. 5e-18;
Matches 112; Conservative 58; Mismatches 148; Indels 85; Gaps 13;

QY 5 WNGSDGPEGAREPPWALPCDERRCSPPFLGA---LVPVTVACLCLFVVGVSGNVVTVM 61
DB 12 WNASASSGSHN--WSLVDPVS-----PMGARAVLPVLYLLVC--TVGLGNTLVY 60

QY 62 LIGRYDRMTTNTNLYLGSMVSDLLILLGLPFDLYRLWRSRPVPGPLCLRLSLYVGE 121
DB 61 VVLYRYAKMTVTNVIINLAVADVLFMLGLPF-LATQNAVSVYWPFGSFLCRLVMTLDGIN 119

QY 122 TYATLLHMTALSVERVLAICRPARVLTTRRVRALIAVLWAVALLSAGPFLFLVGE 181
DB 120 QFTSIFCLMVNSVDYLVVHPLRSARWRPRVAKLASAAVWFSLMSLPLLVFADVQE 179

QY 182 DPGISVVGFLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRCRPSPAQ 241
DB 180 G-----WGTCNLS-----W-----PEPVG 193

QY 242 LGALRWLWTTTAYFFLFLCLSLYGLI-----GRELWSSRRPLRGPAAGRGRGHR 294
DB 194 LWGAAPITTSVLGFFGLPGLVLCVLLIVKVKAAAGMRVGSRR-----RRRSR 243

QY 295 QTVRVLLVVLATIIICWLPFHVGRIIVN-TEDSRMVYFSQYFNIVALQLFYLASINPI 353
DB 244 KVTVMVVVVVLFVGVGWLPPFFIIVNVLNFTLPEEPTSAGLYFFVVVLS--YANSCANPL 301

QY 354 LYNLSIKKYRAAFLKLLARKSRPRGPHRSRDTAGEVAGDTGG 396
DB 302 LYGLFDNFQSRFKALCLR----RGYVEDADAIEPRPKSG 340

RESULT 13
SSRS RAT STANDARD; PRT; 363 AA.
AC P30938;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Somatostatin receptor type 5 (SSSR).
GN SSTR5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=93125499; PubMed=1362243;
RA O'Carroll A.-M., Lolait S.J., Konig M., Mahan L.C.;
```

```
RT RT
RL RL
RN RN
RP RP
RC RC
RX RX
RA RA
RA Panetta R., Greenwood M.T., Warszynska A., Demchyshyn L.L., Day R.,
RA Niznik H.B., Srikant C.B., Patel Y.C.;
RT "Molecular cloning, functional characterization, and chromosomal
RT localization of a human somatostatin receptor (somatostatin receptor
RT type 5) with preferential affinity for somatostatin-28.";
RL Mol. Pharmacol. 45:417-427(1994).
CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PROMINENT IN THE PITUITARY AND SMALL
CC INTESTINE. LOW LEVELS IN ISLETS AND SPLEEN. NOT DETECTED IN
CC KIDNEY, PANCREAS, CEREBELLUM, OR CORTEX.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; L04535; AAA17029.1; -
CC EMBL; U01152; AAC09011.1; -
CC EMBL; X74828; CAA52825.1; -
CC PIR; I57940; I57940.
CC HSP; P02699; I57940.
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPSN.
CC PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.
CC PROSITE; PS00262; G PROTEIN RECEPTOR FL2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family; Lipoprotein; Palmitate
KW DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 36 63 1 (POTENTIAL).
FT DOMAIN 64 73 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 74 99 2 (POTENTIAL).
FT DOMAIN 100 111 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 112 133 3 (POTENTIAL).
FT DOMAIN 134 155 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 156 176 4 (POTENTIAL).
FT DOMAIN 177 196 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 197 221 5 (POTENTIAL).
FT DOMAIN 222 247 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 248 273 6 (POTENTIAL).
FT DOMAIN 274 283 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 284 308 7 (POTENTIAL).
FT DOMAIN 309 363 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 110 185 BY SIMILARITY.
FT LIPID 320 320 PALMITATE (POTENTIAL).
SQ SEQUENCE 363 AA; 39971 MW; 4BD451296061384A CRC64;

Query Match 16.7%; Score 359; DB 1; Length 363;
Best Local Similarity 27.3%; Pred. No. 3e-17;
Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;

QY 5 WNGSDGPEGAREPPWALPCDERRCSPPFLGA---LVPVTVACLCLFVVGVSGNVVTVM 61
DB 12 WNASASSGSHN--WSLVG-----SASPMGARAVLPVLYLLVC--TVGLGNTLVY 60

QY 62 LIGRYDRMTTNTNLYLGSMVSDLLILLGLPFDLYRLWRSRPVPGPLCLRLSLYVGE 121
DB 61 VVLYRYAKMTVTNVIINLAVADVLFMLGLPF-LATQNAVSVYWPFGSFLCRLVMTLDGIN 119
```


Db 61 VVLRHAKMTVTNVIINLAVADVLFMLGLPFLATQNAVVSYPFGSFLCRLVMTLDGIN 120
 QY 122 TYATLLHMTALSVERYLAICRPLARLVTRRRVRLIALVLAVALISAGPFLFLVGVEQ 181
 Db 121 QFTSIFCLMWSVDRIYLVVHPILRSARWRPRVAKASAAVWVFLMSLPLVLFADVQE 180
 QY 182 DPGISVVVGLNGTARIASSPLASSPPLWLRAPPPSPFSGPETAFAAALFSRECRPSPAQ 241
 Db 181 G-----WGICNLSS-----W-----PEPVG 194
 QY 242 LGLRVLMLWTTAYFFLFLCFLSILYGLI-----GRELWSSRRPLRGPAAASRRGRHR 294
 Db 195 LWGAFTTYSVLGFFGLFVLCVLLVVKVKAAGMRVGSRR-----RRRSEP 244
 QY 295 QTVRVLLAVVLAFLICWLPVHGRIIYIN-TEDSRMWYFQYINVAIQLFYLSASINPI 353
 Db 245 KVTSMVVVVVLFVGCWLPFFIVNVMLATLPEEPTSAGLYFFVVVLS--YANSCANPL 302
 QY 354 LYNLSIKKYRAAFAKLLARKSRPRGHRSDTAGEVAGDTGG 396
 Db 303 LYGLSLDNFQSRFRKVLCLR----RGYGMEDADAIEPRPKSG 341

RESULT 14
 SSR3 HUMAN
 ID SSR3 HUMAN STANDARD; PRT; 418 AA.
 AC P32745;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Somatostatin receptor type 3 (SSR3) (SSR-28).
 GN SSR3.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93149123; PubMed=1337145;
 RA Yamada Y., Reisine T., Law S.F., Ihara Y., Kubota A., Kagimoto S.,
 RA Seino M., Seino Y., Bell G.I., Seino S.;
 RT "Somatostatin receptors, an expanding gene family: cloning and
 RT functional characterization of human SSR3, a protein coupled to
 RT adenylyl cyclase.";
 RL Mol. Endocrinol. 6:2136-2142(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93238970; PubMed=8097479;
 RA Corness J.D., Demchyshyn L.L., Seeman P., van Tol H.H.M.,
 RA Srikant C.B., Kent G., Patel Y.C., Niznik H.B.;
 RT "A human somatostatin receptor (SSR3), located on chromosome 22,
 RT displays preferential affinity for somatostatin-14 like peptides.";
 RL FEBS Lett. 321:279-284(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.B., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.B., Graham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,

RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Walli M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuayama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyraud M., Kedra D.,
 RA Seroussi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P., Lane L.,
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tilahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATINS-14 AND -28. THIS RECEPTOR IS
 CC COUPLED VIA PERTUSSIS TOXIN SENSITIVE G PROTEINS TO INHIBITION OF
 CC ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN, PITUITARY AND PANCREAS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M96738; AAA60592.1; -;
 CC EMBL; Z82188; CAB45263.1; -;
 CC PIR; A46226; A46226.
 CC HSSP; P34996; LDDD.
 CC Genew; HGNC:11332; SSTR3.
 CC MIM; 182453; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0004994; F:somatostatin receptor activity; TAS.
 CC GO; GO:0007267; P:cell-cell signaling; TAS.
 CC GO; GO:0007187; P:G-protein signaling; coupled to cyclic nucl. . . ; TAS.
 CC GO; GO:0008628; P:induction of apoptosis by hormones; TAS.
 CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm 1; 1.
 CC PRINTS; PR00237; GPCR_Rhodopsn.
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Multigene family; Polymorphism.
 CC DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 44 69
 CC DOMAIN 70 79 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 80 101 2 (POTENTIAL).
 CC DOMAIN 102 116 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 117 138 3 (POTENTIAL).
 CC DOMAIN 139 161 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 162 181 4 (POTENTIAL).
 CC DOMAIN 182 205 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 206 231 5 (POTENTIAL).
 CC DOMAIN 232 257 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 258 279 6 (POTENTIAL).
 CC DOMAIN 280 293 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 294 316 7 (POTENTIAL).
FT DOMAIN 317 418 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 116 191 BY SIMILARITY.
FT DOMAIN 346 360 GLU-RICH (ACIDIC).
FT VARIANT 411 411 S->T (IN dSNP:229568).
FT /FTID=VAR.011853.
SQ SEQUENCE 418 AA; 45847 MW; 1227095F801190C4 CRC64;
Query Match 16.6%; Score 357.5; DB 1; Length 418;
Best Local Similarity 26.9%; Pred. No. 4.2e-17;
Matches 101; Conservative 63; Mismatches 146; Indels 65; Gaps 10;
QY 11 PEGAREPPWALPCDERRCSPPFL-----GALVPVAVCLCLFVVGVSQNVVTVMLIGR 65
DB 15 PENA-SSAWPPDATLGNVSAGSPAGLAVSGVLPLVVLVC--VGLLGNLSLVIVVLR 71
QY 66 YRDMRTTNLYLGSMAVSDLLILGLPDLRLWRSRPWVFGPLLCRLSLYVGEQTYAT 125
DB 72 HTASPSVTNVIYNLALADELMGLGFP-LAAQNALSYWPFGLMCLVMAVDGINQFTS 130
QY 126 LLHMTALSVRYLAICRPLARVLVTRRRVRLAVLMAVALLSAGPFLVGVGEQDPI 185
DB 131 IFCLTVMSVDRLAVVHTPSARWRTAPVARTVSAVAVVAVVAVVLPVVVFSGV- 184
QY 186 SVVPLGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAFAALFSRECR---PSPAOL 242
DB 185 -----PRGMST-----CHMQWPEPAAA 201
QY 243 GALKRVLWVTVTAFFLPFLCLSLYGLIGRELWSSRLRGPAAAGRGRHROTVRVLLV 302
DB 202 WRAGFIITTAALGFGLPFLVLCYLLIVVKSAGRRVWAPSQRERRSERRVTRVVA 261
QY 303 VVLAFLICWLPVHGRVRIIYNTE-DSRMVFSQVFNVALQLFVLSINPILYNLSKK 361
DB 262 VVALFVLCWMPFFVNLVNVVCPLEPAPFGLVFLVALP--YANSCANPILYGFUSYR 319
QY 362 YRAAFKLLARKSR 376
DB 320 FK-GFRVLLRPSR 333

RESULT 15

GP39 HUMAN
ID GP39 HUMAN STANDARD; PRT; 453 AA.
AC O43194;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative G protein-coupled receptor GPR39.
GN GPR39.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98110578; PubMed=9441746;
RA McKee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,
RA Hrenuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
RT "Cloning and characterization of two human G protein-coupled receptor
RT genes (GPR38 and GPR39) related to the growth hormone secretagogue
RT and neurotensin receptors.";
RL Genomics 46:426-434(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.emb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF034633; AAC26082.1; -
DR Genew; HGNC:4496; GPR39.
DR MIM; 602886; -
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0007186; P:G-protein coupled receptor protein signal. . .; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 26 46 1 (POTENTIAL).
FT DOMAIN 47 71 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 72 92 2 (POTENTIAL).
FT DOMAIN 93 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 131 3 (POTENTIAL).
FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 4 (POTENTIAL).
FT TRANSMEM 176 223 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 224 243 5 (POTENTIAL).
FT DOMAIN 244 280 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 281 301 6 (POTENTIAL).
FT DOMAIN 302 321 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 322 347 7 (POTENTIAL).
FT TRANSMEM 348 428 CYTOPLASMIC (POTENTIAL).
FT DISULFID 108 210 BY SIMILARITY.
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 453 AA; 51328 MW; 8E3A233420D9021E CRC64;

Query Match 16.5%; Score 355.5; DB 1; Length 453;

Best Local Similarity 28.5%; Pred. No. 6.2e-17;

Matches 111; Conservative 65; Mismatches 169; Indels 45; Gaps 13;

QY 20 PALPPCDERRCS-----PPFLGALVPVT--AVCLCLFVVGVSQNVTV---MLIGR 65

DB 4 PSLPGSD---CSQIDHSHVPEFVATWIKITLILVLIIFVMLGLGNSATIRVTVLQK 60

QY 66 YRDMRTTNLYLGSMAVSDLLI-LLGLPDLVLR-LWRSRWVFGPLLCRLSLYVGEQTY 123

DB 61 KGYLQKEVTDHVMVSLACSDILVELIGHPMEFYIWNPLTTSYTLCKLHTEFLPEACSY 120

QY 124 ATLLHMTALSVRYLAICRPLARVLVTRRRVRLAVLMAVALLSAGPFLVGVGEQDP 183

DB 121 ATLLHVTLSFERYIAICHPRYKAVSGPCQVKLLIGFVWVTSALVALPLLPANGTEY-P 179

QY 184 GISVVGGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAFAALFSRECPSPAQ-- 241

DB 180 LVN-VPSHRGLTCNRSSTRHEQ-----PETSMSICTNLSSRWTVFQSS 223

QY 242 -LGALRVMLWVTVTAFFLPFLCLSLYGLIGRELWSSRRP--LRGPAASGRGRHROTVR 298

DB 224 IFGAFVYLVLLVSVAFCMNMVQVLMKSKGSLAGTRPPQKRKSESESRRTARRTII 283

QY 299 VLVVAVLAFIICWLPFHVGRVRIIYI---NTEDSRMMYFSQYFNIVALQ--LFYLSASINPIL 354

DB 284 FLALIVTVLAVCMVNPQIRRIAMAAKPKHDWTSYFRAYMILLPFPSTFFYLSSVINPLL 343

QY 355 YNLSIKKYRAAAPKLLARKSRGRGFRHSR 384

DB 344 YTVSSQOFRFVQVQLCCRLSLOHANHEKR 373

Search completed: January 1, 2004, 06:23:01

Job time : 16.4561 secs

THIS PAGE BLANK (USPTO)